PCT/EP 98/03318

09/445201

Europäisches Patentamt

European Patent Office

Office européen des brevets

REC'D 21 AUG 1998
WIPO PCT

Bescheinigung

Certificate

Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

97108959.4

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

M.B. RIJLING

DEN HAAG, DEN
THE HAGUE, 13/08/98
LA HAYE, LE

Europäisches Patentamt | European Patent Office | Office européen des brevets

Blatt 2 der Bescheinigung
Sheet 2 of the certificate
Page 2 de l'attestation

Anmeldung Nr.:
Application no.: 97108959.4
Demande n°:

Anmeldetag:
Date of filing: 03/06/97
Date de dépôt:

Anmelder:
Applicant(s):
Demandeur(s):
Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V.
Berlin
GERMANY

Bezeichnung der Erfindung:
Title of the invention:
Titre de l'invention:
Regulatory sequences capable of conferring expression of a heterologous DNA sequence in endothelial cells in vivo and uses thereof

In Anspruch genommene Prioriät(en) / Priority(ies) claimed / Priorité(s) revendiquée(s)

Staat:
State:
Pays:

Tag:
Date:
Date:

Aktenzeichen:
File no.
Numéro de dépôt:

Internationale Patentklassifikation:
International Patent classification:
Classification internationale des brevets:
/

Am Anmeldetag benannte Vertragstaaten:
Contracting states designated at date of filing: AT/BE/CH/DE/DK/ES/FI/FR/GB/GR/IE/IT/LI/LU/MC/NL/PT/SE
Etats contractants désignés lors du depôt:

Bemerkungen:
Remarks:
Remarques:

Our Ref.: B 1730 EP

REGULATORY SEQUENCES CAPABLE OF CONFERRING EXPRESSION OF A HETEROLOGOUS DNA SEQUENCE IN ENDOTHELIAL CELLS IN VIVO AND USES THEREOF

The present invention relates to recombinant DNA molecules comprising the regulatory sequence(s) of an intron of the Vascular Endothelial Growth Factor (VEGF) receptor-2 gene (Flk-1) or of a gene homologous to the Flk-1 gene, being capable of conferring expression of a heterologous DNA sequence in endothelial cells in vivo. The present invention also provides vectors comprising said recombinant DNA molecules. The present invention additionally relates to pharmaceutical and diagnostic compositions comprising such recombinant DNA molecules and vectors. Furthermore, the present invention relates to cells and transgenic non-human animals, comprising the aforementioned recombinant DNA molecules or vectors stably integrated into their genome and their use for the identification of substances capable of suppressing or activating transcription of a gene in endothelial cells. The present invention also relates to the use of the before described recombinant DNA molecules and vectors for the preparation of pharmaceutical compositions for treating, preventing, and/or delaying a vascular or tumorous disease in a subject. Furthermore, the recombinant DNA molecules and vectors of the invention can be used for the preparation of pharmaceutical compositions for inducing a vascular or tumorous disease in a non-human animal.

In the field of neuroscience and medical therapy, there is a great demand for test systems to study the function and interaction of gene products, the malfunction or expression of which cause vascular and/or tumorous diseases. Such systems would also be suitable for drug development against such diseases. A prominent example

for gene products involved in vascular diseases are angiogenic growth factors and their endothelial receptors which play a major role in the formation of the embryonic vascular system and in certain angiogenesis-dependent diseases, such as solid tumor growth or retinopathy. The Kinase-insert Domain-containing Receptor/fetal liver kinase-1 (KDR/Flk-1) in the following referred to as Flk-1 and Flt-1 are high affinity signaling receptors for the endothelial mitogen, vascular endothelial growth factor (VEGF) (Connolly, J. Clin. Invest. 84 (1989), 1470-1478; Leung, Science 246 (1989), 1306-1309). Through interactions with its receptors, VEGF plays critical roles in growth and maintenance of vascular endothelial cells and in the development of new blood vessels in physiologic and pathologic states (Aiello, New Engl. J. Med. 331 (1994), 1480-1487; Shweiki, Nature 359 (1992), 843-845; Berkman, J. Clin. Invest. 91 (1993), 153-159). The patterns of embryonic expression of VEGF suggest that it is crucial for differentiation of endothelial cells from hemangioblasts and for development of blood vessels at all stages of growth (Jakeman, Endocrinology 133 (1993), 848-859; Breier, Development 114 (1992), 521-532). Among many potentially angiogenic factors, VEGF is the only one with patterns of expression, secretion, and activity that suggest a specific angiogenic function in normal development (Klagsbrun, Current Biology 3 (1993), 699-702).

High-affinity receptors for VEGF are found only on endothelial cells, and VEGF binding has been demonstrated on macro- and microvascular endothelial cells and in quiescent and proliferating endothelial cells (Jakeman, Endocrinology 133 (1993), 848-859; Jakeman, Clin. Invest. 89 (1992), 244-253). The Flk-1 and Flt-1 have been identified as candidate VEGF receptors by affinity cross-linking and competition-binding assays (de Vries, Science 255 (1992), 989-991; Millauer, Cell 72 (1993), 835-846; Terman, Biochem. Biophys. Res. Commun. 187 (1992), 1579-1586). These two receptor tyrosine kinases contain seven similar extracellular immunoglobulin domains and a conserved intracellular tyrosine kinase domain interrupted by a kinase insert (de Vries, Science 255 (1992), 989-991; Matthews, Proc. Natl. Acad. Sci. U.S.A 88 (1991), 9026-9030; Terman, Oncogene 6 (19..), 1677-1683); they are expressed specifically by endothelial cells *in vivo* (Millauer, Cell 72 (1993), 835-846; Peters, Proc. Natl. Acad. Sci. USA 90 (1993), 7533-7537;

Yamaguchi, Development 118 (1993), 489-498). *In situ* hybridization in the developing mouse has demonstrated that Flk-1 is expressed in endothelial cells at all stages of development, as well as in the blood island in which endothelial cell precursors first appear (Millauer, Cell 72 (1993), 835-846). Flk-1 is a marker for endothelial cell precursors at their earliest stages of development (Yamaguchi, Development 118 (1993), 489-498).

The vascular endothelium is critical for physiologic responses including thrombosis and thrombolysis, lymphocyte and macrophage homing, modulation of the immune response, and regulation of vascular tone. The endothelium is also intimately involved in the pathogenesis of vascular diseases such as atherosclerosis (Ross, Nature 362 (1993), 801-809). Although a number of genes expressed in the endothelium have been characterized (Collins, J. Biol. Chem. 266 (1991), 2466-2473; Iademarco, J. Biol. Chem. 267 (1992), 16323-16329; Jahroudi, Mol. Cell. Biol. 14 (1994), 999-1008; Lee, J. Biol. Chem. 265 (1990), 10446-10450), expression of these genes is either not limited to vascular endothelium (e.g., the genes encoding von Willebrand factor, endothelin-1, vascular cell adhesion molecule-1), or is restricted to specific subpopulations of endothelial cells (e.g., the gene for endothelial-leukocyte adhesion molecule-1). Flk-1 (also known as VEGF-receptor 2) is expressed in endothelial cells during embryonic and postnatal development. The Flk-1 receptor is the first endothelial receptor to be expressed in endothelial cell precursors during embryonic vascular development. Gene targeting experiments in transgenic mice have demonstrated that this receptor is essential for endothelial cell differentiation (Shalaby, Nature 376 (1995), 62-66). Furthermore, in a variety of tumors, Flk-1 receptor expression is re-induced in the tumor vasculature, and it has been shown that signaling via the Flk-1 receptor is required for tumor vascularization and growth (Millauer, Nature 367 (1994), 576-579).

Recently, in vitro studies with the upstream region of the human Flk-1 gene (Patterson, J. Biol. Chem. (1995), 23111-23118) showed that DNA fragments located in the 5' flanking region of the human Flk-1 gene mediate expression of a reporter gene.

For studying all aspects of genes involved in vascular diseases such as atherosclerosis, however, the system described by Patterson, supra, suffers from several drawbacks. For example, promoter activity of the 5'-flanking region used by Patterson was also observed in cell types which do not express the Flk-1 gene naturally. Furthermore, the promoter fragment employed in Patterson, supra, was not shown to be expressed in vivo in its natural background. In order to specifically suppress or confer endothelium specific gene expression and for the development of endothelium specific drugs, however, one needs test systems which closely resemble the regulation of the Flk-1 expression in vivo since otherwise non-informative or even false positive results may be obtained.

Thus, the technical problem of the present invention is to provide a means that allows studying the expression of nucleic acids specifically in endothelial cells in vivo, preferably at all stages of development.
The solution to this technical problem is achieved by providing the embodiments characterized in the claims.

Accordingly, the invention relates to a recombinant DNA molecule comprising:

(a) a first regulatory sequence of an intron of the Vascular Endothelial Growth factor (VEGF) receptor-2 (Flk-1) gene or of a gene homologous to the Flk-1 gene being capable of conferring expression in endothelial cells in vivo; and
(b) operatively linked thereto a heterologous DNA sequence.

In accordance with the present invention, a regulatory sequence driving the expression of a heterologous DNA sequence in substantially all endothelial cells in vivo, preferably at substantially all stages of development has been identified. Said regulatory sequence is suitable to direct the expression of a heterologous DNA sequence in the above-mentioned cells. The recombinant DNA molecule of the invention allows studying the function and interaction of proteins which are expressed in the endothelium of, for example, humans and the malfunction, and/or unregulated expression of which is supposed to be the or a causative agent of

vascular and tumorous diseases. Thus, the regulatory sequences of the invention are particularly suited and useful for the engineering of transgenic cells and non-human animals which can serve as a test system for the development of drugs for the treatment of vascular and tumorous diseases of endothelial origin.

In the context of the present invention the term "a first regulatory sequence of an intron of the Flk-1 gene" means a nucleotide sequence of the first intron of the murine Flk-1 gene including the regulatory sequences which are capable of conferring the specific expression of a heterologous DNA sequence in endothelial cells, preferably at all stages of development.

In accordance with the present invention, it was shown for the first time that the expression of a heterologous DNA sequence under the control of said regulatory sequence occurs specifically in endothelial cells at substantially all stages of development in a mouse. In order to characterize cis-acting regulatory sequences contained in the Flk-1 gene, recombinant bacteriophage lambda clones containing mouse genomic DNA (Mouse strain 129/Sv) have been isolated encompassing a 21 kilo base pair (kb) region of the mouse Flk-1 gene, contained in the DNA insertions of λ phages 6 and 16, including approximately 15 kb of 5' flanking sequences, exons 1, 2 and 3, and introns 1 and 2 (Fig. 4A). The DNA sequence of a 12.8 kb region spanning from position -6.65 kb (the affixes - and + refer to the nucleotide position relative to the transcriptional start site as shown in Fig. 1 which corresponds to nucleotide position 6661 of SEQ ID NO: 1) to position +6.15 (located in the third exon) was determined (SEQ ID NO: 1). Reporter gene studies were performed in order to characterize regulatory cis-acting elements of the Flk-1 gene. Initial studies focused on the role of 5' flanking sequences of the Flk-1 gene ("Flk-1 promoter") in mediating endothelium-specific expression in cultured bovine aortic endothelial (BAE) cells (Rönicke, Circulation Research 79 (1996), 277-285). In these studies, it was found that a promoter fragment ranging from -624 to +299 mediated high expression of the luciferase reporter gene following transient transfection in BAE cells. Experiments with transgenic mouse embryos performed in accordance with the present invention revealed, however, that the murine promoter DNA fragments were

not sufficient to mediate endothelium-specific reporter gene expression in vivo. Surprisingly, however, when a Flk-1 promoter fragment (ranging from, e.g., -624 to +299 bp) was combined with a 2.3 kb fragment of the first Flk-1 intron, endothelium specific expression of a lacZ reporter gene in mouse embryos was obtained. Thus, the first intron (nucleotides 7027 to 10642 of SEQ ID NO: 1) of the mouse Flk-1 gene is essential for endothelium specific gene expression. This is a novel finding because the sequences described in previous publications (Patterson, supra; Rönicke, supra) are, in contrast to the expectations and interpretations of the prior art, not sufficient to mediate endothelium-specific expression in vivo. These results obtained in accordance with the present invention demonstrate that the regulatory sequences located in the intron of the Flk-1 gene mediating endothelium-specific expression can be used to direct expression of heterologous genes in the vasculature.

The genomic DNA of the murine Flk-1 gene comprising the intron regulatory sequences can be obtained from liver of mouse strain 129/SV, or, for example, by screening a phage library of liver genomic DNA in the vector λFixII (Stratagene, La Jolla, CA) generated by conventional methods known in the art.

The term "regulatory sequence of a gene homologous to the Flk-1 gene" also includes promoter regions and regulatory sequences of a gene from another species, for example, humans and other mammals which is homologous to the Flk-1 gene of mouse and which confers the same expression pattern. Such regulatory sequences are characterized by their capability of conferring expression of a heterologous DNA sequence specifically in endothelial cells in vivo, preferably at all stages of development. Thus, according to the present invention, regulatory sequences from other species can be used that are functionally homologous to the regulatory sequences of the intron of the Flk-1 gene from mouse, or regulatory sequences of genes that display an identical pattern of expression, in the sense of being expressed in the endothelium, preferably at all stages of development.

It is possible for the person skilled in the art to isolate by employing the known Flk-1 gene from mouse, corresponding genes from other species, for example, humans

and other mammals. This can be done by conventional techniques known in the art, for example, by using Flk-1 gene sequences as a hybridization probe or by designing appropriate PCR primers. It is then possible to isolate the corresponding regulatory sequences by conventional techniques and test them for their expression pattern. For this purpose, it is, for instance, possible to fuse the regulatory sequences to a reporter gene, such as the luciferase or green fluorescent protein (GFP) encoding genes and assess the expression of the reporter gene in transgenic animals, for example in mice. The partial nucleotide sequence of the human Flk-1 gene may be obtained from Genbank Acc. No. X89776; Patterson, supra; Terman, Bicochem. Biophys. Res. Comm. 187 (1992), 1579-1586; Genbank Acc. No. X61656.

The present invention also relates to recombinant DNA molecules comprising regulatory sequences which are substantially identical to that of the Flk-1 intron or to an intron of a homologous gene or to fragments thereof and which are able to confer specific expression in endothelial cells, preferably at all stages of development in mouse or other mammals.

Such regulatory sequences differ at one or more positions from the above-mentioned regulatory sequences but still have the same specificity, namely they comprise the same or similar sequence motifs responsible for the above described expression pattern. Preferably such regulatory sequences hybridize to one of the above-mentioned regulatory sequences, most preferably under stringent conditions. Particularly preferred are regulatory sequences which share at least 85%, more preferably 90-95%, and most preferably 96-99% sequence identity with one of the above-mentioned regulatory sequences and have the same specificity. Such regulatory sequences also comprise those which are analogues or derivatives, for example by way of nucleotide deletion(s), insertion(s), substitution(s), addition(s), and/or recombination(s) and/or any other modification(s) known in the art either alone or in combination in comparison to the above-described nucleotide sequence. Methods for introducing such modifications in the nucleotide sequence of the regulatory sequences of the invention are well known to the person skilled in the art and described, for example, in Sambrook (Molecular cloning; A Laboratory Manual,

Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY (1989)). All such fragments, analogues and derivatives of the regulatory sequence of the invention are included within the scope of the present invention, as long as the essential characteristic regulatory properties as defined above remain unaffected in kind. It is also immediately evident to the person skilled in the art that further regulatory sequences may be added to the regulatory sequences of the invention. For example promoters, transcriptional enhancers and/or sequences which allow for induced expression of the regulatory sequences of the invention may be employed. A suitable inducible system is for example tetracycline-regulated gene expression which is described by, e.g., Gossen (Proc. Natl. Acad. Sci. USA 89 (1992), 5547-5551; Trends Biotech. 12 (1994), 58-62).
The expression conferred by the regulatory sequences of the invention may not be exclusively limited to the above-described specificity but may also occur in, e.g., neuronal cells, including neural retinal progenitor cells at all or different stages of development and haematopoietic cells (Yang, J. Neurosci. 16 (1996), 6089-6099).

The term "further regulatory sequences" refers to sequences which influence the specificity and/or level of expression, for example in the sense that they confer cell and/or tissue specificity or developmentally and/or inducible regulated gene expression. Such regions can be located upstream of or comprising the transcription initiation site, such as a promoter, but can also be located downstream of it, e.g., in transcribed but not translated leader sequences.
The term "promoter" refers to the nucleotide sequences necessary for transcription initiation, i.e. RNA polymerase binding, and also includes, for example, the TATA box.
The term "in vivo" for the purpose of the present invention is used for cells in an organism as opposed to cells growing in culture (in vitro).
The term "heterologous" with respect to the DNA sequence being operatively linked to the promoter of the invention means that said DNA sequence is not naturally linked to the regulatory sequences comprised in the recombinant DNA molecule of the invention.

In a preferred embodiment said first regulatory sequence of the invention comprises a GATA-binding site, an AP-1 binding site, an SP1 binding site, a PEA3 consensus sequence or any combination(s) thereof. A functional analysis of the first 6.5 kbp of the transcribed region of the murine Flk-1 genes lead to the identification of a endothelial-specific positive regulatory element. This regulatory sequence is located in the region between the XhoI and BamHI restriction site in the first intron of the Flk-1 gene (cf. Fig. 4A). It is functional in both orientations since the intron enhancer was used in an antiparallel manner with respect to the Flk-1 promoter fragment in the construct referred to as 3'-In 1; see Example 2 hereinbelow. A sequence analysis of the intron lead to the identification of two potential GATA binding sites (+1927 Bp, +3514 Bp); a potential AP-1 binding site (+2210 Bp) and two PEA3 consensus sequences (+3494 Bp, +3741 Bp); see Fig. 1.

Preferably, said first regulatory sequence is selected from the group consisting of

(a) DNA sequences comprising a nucleotide sequence as given in SEQ ID NO: 1, preferably from nucleotide 8260 to nucleotide 10560;
(b) DNA sequences comprising the nucleotide sequence of the human Flk-1-intron;
(c) DNA sequences comprising a nucleotide sequence which hybridizes with a nucleotide sequence of (a) or (b) under stringent conditions;
(d) DNA sequences comprising a nucleotide sequence which is conserved in the nucleotide sequences of (a) and (b); and
(e) DNA sequences comprising a fragment, analogue or derivative of a nucleotide sequence of any one of (a) to (d) capable of conferring expression in endothelial cells.

In a particularly preferred embodiment of the present invention, the regulatory sequences comprise the nucleotides 8260 to 10560 and most preferably nucleotides 8560 to 10400 **[bitte ggf. weiter eingrenzen]** of the nucleotide sequence as set

forth in SEQ ID No. 1 or a fragment thereof, which still confers expression in endothelial cells, preferably at all stages of development.

In a further preferred embodiment of the invention the heterologous DNA sequence of the recombinant DNA molecules described above is operatively linked to further regulatory sequences. Expression comprises transcription of the nucleic acid molecule, preferably into a translatable mRNA. Regulatory elements ensuring expression in eukaryotic cells, preferably mammalian cells, are well known to those skilled in the art. They normally comprise promoters ensuring initiation of transcription and optionally poly-A signals ensuring termination of transcription and stabilization of the transcript. Additional regulatory elements may include transcriptional as well as translational enhancers. Preferably said further regulatory sequence is a promoter and/or a 3'-untranslated region.
Although some endothelial-specific promoters have been characterized, e. g. of the genes for von Willebrand factor (Jahroudi, Mol. Cell Biol. 14 (1994), 999-1008), Endothelin-1 (Lee, J. Biol. Chem. 265 (1990), 10446-10450), E-selectin (Collins, J. Biol. Chem. 266 (1991), 2466-2473), Tie-2 (Schlaeger, Development 121 (1995), 1089-1098), VCAM-1 (Iademarco, J. Biol. Chem. 267 (1992), 16323-16329) and endothelial NO-synthase (Zhang, J. Biol. Chem. 270 (1995), 15320-15326) these genes are neither specific for proliferating endothelium, nor necessary for endothelial cell determination. Due to the present invention these promoters can now be combined with the regulatory sequences of the invention in order to mediate endothelium specific gene expression of heterologous DNA sequences.

In a preferred embodiment the above mentioned promoter is a promoter of hypoxia inducible genes, genes encoding growth factors such as VEGF, PDGF or Fibroblast growth factor or their receptors or glycolytic enzymes.

In a particularly preferred embodiment said promoter comprises a DNA sequence selected from the group consisting of

(a) DNA sequences comprising the nucleotide sequence as given in SEQ ID NO:1 from nucleotide 6036 to nucleotide 6959;

(b) DNA sequences comprising the nucleotide sequence of the human Flk-1 promoter;

(c) DNA sequences comprising a nucleotide sequence which hybridizes with a nucleotide sequence of (a) or (b) under stringent conditions;

(d) DNA sequences comprising a nucleotide sequence which is conserved in the nucleotide sequences of (a) and (b); and

(e) DNA sequences comprising a fragment, analogue or derivative of a nucleotide sequence of any one of (a) to (d).

At least one of the aforedescribed DNA sequences may be preferably of human or murine origin although other sources may be employed as well.

Preferably, the heterologous DNA sequence being operatively linked to the regulatory sequences is located 5' to the regulatory sequence of the invention.

In a further preferred embodiment, the heterologous DNA sequence of the above-described recombinant DNA molecules encodes a peptide, protein, antisense RNA, sense RNA and/or ribozyme. The recombinant DNA molecule or vector of the invention can be used alone or as part of a vector to express heterologous DNA sequences, which, e.g., encode proteins other than Flk-1, in cells of the blood vessel wall, i.e., endothelial cells, for, e.g., gene therapy or diagnostics of vascular diseases such as atherosclerosis. The recombinant DNA molecule or vector containing DNA sequence encoding a protein of interest is introduced into endothelial cells which in turn produce the protein of interest. For example, sequences encoding t-PA (Pennica, Nature 301 (1982), 214), p21 cell cycle inhibitor (El-Deiry, Cell 75 (1993), 817-823), or nitric oxide synthase (Bredt, Nature 347 (1990), 768-770) may be operatively linked to the endothelial cell-specific regulatory sequences of the invention and expressed in endothelial cells. For example, thrombolytic agents can be expressed under the control of the endothelial cell-specific regulatory sequences

of the invention for expression by vascular endothelial cells in blood vessels, e.g., vessels occluded by aberrant blood clots. Other heterologous proteins, e.g., proteins which inhibit smooth muscle cell proliferation, e.g., interferon-γ and atrial natriuretic polypeptide, may be specifically expressed in endothelial cells to ensure the delivery of these therapeutic peptides to an atherosclerotic lesion or an area at risk of developing an atherosclerotic lesion, e.g., an injured blood vessel.

The endothelial cell-specific regulatory sequences of the invention may also be used in gene therapy to promote angiogenesis to treat diseases such as peripheral vascular disease or coronary artery disease (Isner, Circulation 91 (1995), 2687-2692). For example, the regulatory sequences of the invention can be operatively linked to sequences encoding cellular growth factors which promote angiogenesis, e.g., VEGF, acidic fibroblast growth factor, basic fibroblast growth factor and the like.

In a most preferred embodiment of the present invention, said protein is selected from the group consisting of Vascular Endothelial Growth Factor (VEGF), Hypoxia Inducible Factors (HIF), HIF-Related Factor (HRF), tissue plasminogen activator, p21 cell cycle inhibitor, nitric oxide synthase, interferon-γ, atrial natriuretic polypeptide and monocyte chemotactic proteins.

In another particularly preferred embodiment of the invention, said protein is a scorable marker, preferably luciferase, green fluorescent protein or lacZ. This embodiment is particularly useful for simple and rapid screening methods for compounds and substances described herein below capable of modulating the expression of genes in the endothelium. For example, endothelial cells can be cultured with VEGF in the presence and absence of the candidate compound in order to determine whether the compound affects the expression of genes which are under the control of regulatory sequences of the invention, which can be measured, e.g., by monitoring the expression of the above-mentioned marker. It is also immediately evident to those skilled in the art that other marker genes may be employed as well, encoding, for example, selectable marker which provide for the direct selection of compounds which induce or inhibit the expression of said marker.

The regulatory sequences of the invention may also be used in methods of antisense therapy. Antisense therapy may be carried out by administering to an animal or a human patient, a recombinant DNA containing the endothelial cell-specific regulatory sequences of the invention operably linked to a DNA sequence, i.e., an antisense template which is transcribed into an antisense RNA. The antisense RNA may be a short (generally at least 10, preferably at least 14 nucleotides, and up to 100 or more nucleotides) nucleotide sequence formulated to be complementary to a portion of a specific mRNA sequence. Standard methods relating to antisense technology have been described (Melani, Cancer Res. 51 (1991), 2897-2901). Following transcription of the DNA sequence into antisense RNA, the antisense RNA binds to its target mRNA molecules within a cell, thereby inhibiting translation of the mRNA and down-regulating expression of the protein encoded by the mRNA. For example, an antisense sequence complementary to a portion of or all of the Flk-1 (KDR) mRNA (Terman, Oncogene 6 (1991), 1677-1683 and Terman (1992), supra) would inhibit the expression of Flk-1, which in turn would inhibit angiogenesis. Such antisense therapy may be used to treat cancer, particularly to inhibit angiogenesis at the site of a solid tumor, as well as other pathogenic conditions which are caused by or exacerbated by angiogenesis, e.g., inflammatory diseases such as rheumatoid arthritis, and diabetic retinopathy.

The expression of other endothelial cell proteins may also be inhibited in a similar manner, for example, endothelial cell proteins such as cell cycle proteins (thereby inhibiting endothelial cell proliferation, and therefore, angiogenesis); coagulation factors such as von Willebrand factor; and endothelial cell adhesion factors, such as ICAM-1 and VCAM-1 (Bennett, J. Immunol. 152 (1994), 3530-3540).

Thus, in a further preferred embodiment of the present invention, said antisense RNA or said ribozyme are directed against a gene involved in vasculogenesis and/or angiogenesis and/or tumors of endothelial origin.

In a further embodiment, the invention relates to nucleic acid molecules of at least 15 nucleotides in length hybridizing specifically with a regulatory sequence as

described above or with a complementary strand thereof. This means that they hybridize, preferably under stringent conditions, specifically with the nucleotide sequences as described above and show no or very little cross-hybridization with nucleotide sequences having no or substantially different regulatory properties. Such nucleic acid molecules may be used as probes and/or for the control of gene expression. Nucleic acid probe technology is well known to those skilled in the art who will readily appreciate that such probes may vary in length. Preferred are nucleic acid probes of 17, 18, 19, 20 to 25 and 25 to 35 nucleotides in length. Of course, it may also be appropriate to use nucleic acids of up to 100 and more nucleotides in length. The nucleic acid probes of the invention are useful for various applications. On the one hand, they may be used as PCR primers for amplification of regulatory sequences according to the invention. Another application is the use as a hybridization probe to identify regulatory sequences hybridizing to the regulatory sequences of the invention by homology screening of genomic DNA libraries. Nucleic acid molecules according to this preferred embodiment of the invention which are complementary to a regulatory sequence as described above may also be used for repression of expression of a gene comprising such regulatory sequences, for example due to an antisense or triple helix effect or for the construction of appropriate ribozymes (see, e.g., EP-B1 0 291 533, EP-A1 0 321 201, EP-A2 0 360 257) which specifically cleave the (pre)-mRNA of a gene comprising a regulatory sequence of the invention. Selection of appropriate target sites and corresponding ribozymes can be done as described for example in Steinecke, Ribozymes, Methods in Cell Biology 50, Galbraith et al. eds Academic Press, Inc. (1995), 449-460. Furthermore, the person skilled in the art is well aware that it is also possible to label such a nucleic acid probe with an appropriate marker for specific applications, such as for the detection of the presence of a nucleic acid molecule of the invention in a sample derived from an organism.

Such molecules may either be DNA or RNA or a hybrid thereof. Furthermore, said nucleic acid molecule may contain, for example, thioester bonds and/or nucleotides analogues, commonly used in oligonucleotide anti-sense approaches. Said modifications may be useful for the stabilization of the nucleic acid molecule against

endo- and/or exonucleases in the cell. Said nucleic acid molecules may also be transcribed by an appropriate vector containing a chimeric gene which allows for the transcription of said nucleic acid molecule in the cell. Such nucleic acid molecules may further contain ribozyme sequences which specifically cleave the (pre)-mRNA comprising the regulatory sequence of the invention. Furthermore, oligonucleotides can be designed which are complementary to a regulatory sequence of the invention (triple helix; see Lee, Nucl. Acids Res. 6 (1979), 3073; Cooney, Science 241 (1988), 456 and Dervan, Science 251 (1991), 1360), thereby preventing transcription and the production of the encoded protein.
The present invention also relates to vectors, particularly plasmids, cosmids, viruses, bacteriophages used conventionally in genetic engineering that comprise a recombinant DNA molecule of the invention. Preferably, said vector is an expression vector and/or a targeting vector. Expression vectors derived from viruses such as retroviruses, vaccinia virus, adeno-associated virus, herpes viruses, or bovine papilloma virus, may be used for delivery of the recombinant DNA molecule or vector of the invention into targeted cell population. Methods which are well known to those skilled in the art can be used to construct recombinant viral vectors; see, for example, the techniques described in Sambrook, Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory (1989) N.Y. and Ausubel, Current Protocols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y. (1989). Alternatively, the recombinant DNA molecules and vectors of the invention can be reconstituted into liposomes for delivery to target cells.

The present invention furthermore relates to host cells transformed with a DNA molecule or vector of the invention. Said host cell may be a prokaryotic or eukaryotic cell. The vector or recombinant DNA molecule of the invention which is present in the host cell may either be integrated into the genome of the host cell or it may be maintained extrachromosomally. In this respect, it is also to be understood that the recombinant DNA molecule of the invention can be used for "gene targeting" and/or "gene replacement", for restoring a mutant gene or for creating a mutant gene via homologous recombination.

The host cell can be any prokaryotic or eukaryotic cell, such as a bacterial, insect, fungal, plant or animal cell. Preferred fungal cells are, for example, those of the genus Saccharomyces, in particular those of the species S. cerevisiae. Suitable mammalian cell lines comprise Saos-2 human osteosarcoma cells (ATCC HTB-85), HeLa human epidermoid carcinoma cells (ATCC CRL-7923), HepG2 human hepatoma cells (ATCC HB-8065), human fibroblasts (ATCC CRL-1634), U937 human histiocytic lymphoma cells (ATCC CRL-7939), RD human embryonal rhabdomyosarcoma cells (ATCC CCL-136), MCF7 human breast adenocarcinoma cells (ATCC HTB-22), JEG-3 human choriocarcinoma cells (ATCC HB36), A7r5 fetal rat aortic smooth muscle cells (ATCC CRL-1444), and NIH 3T3 mouse fibroblasts (ATCC CRL-1658) obtainable from the American Type Culture Collection. Primary-culture HUVEC may be obtained from Clonetics Corp. (San Diego, CA) and can be grown in EGM medium containing 2% fetal calf serum (Clonetics). Primary-culture human aortic and intestinal smooth muscle cells can also be obtained from Clonetics Corp. Most preferably said host cell is an endothelial cell or derived therefrom, such as BAE cells.

Moreover, the present invention relates to a pharmaceutical composition comprising at least one of the aforementioned recombinant DNA molecules or vectors of the invention, either alone or in combination, and optionally a pharmaceutically acceptable carrier or excipient. Examples of suitable pharmaceutical carriers are well known in the art and include phosphate buffered saline solutions, water, emulsions, such as oil/water emulsions, various types of wetting agents, sterile solutions etc. Compositions comprising such carriers can be formulated by well known conventional methods. These pharmaceutical compositions can be administered to the subject at a suitable dose. Administration of the suitable compositions may be effected by different ways, e.g. by intravenous, intraperitoneal, subcutaneous, intramuscular, topical or intradermal administration. The dosage regimen will be determined by the attending physician and other clinical factors. As is well known in the medical arts, dosages for any one patient depends upon many factors, including the patient's size, body surface area, age, the particular compound to be

administered, sex, time and route of administration, general health, and other drugs being administered concurrently. Dosages will vary but a preferred dosage for intravenous administration of DNA is from approximately 10^6 to 10^{22} copies of the DNA molecule. The compositions of the invention may be administered locally or systemically. Administration will generally be parenterally, e.g., intravenously; DNA may also be administered directly to the target site, e.g., by biolistic delivery to an internal or external target site or by catheter to a site in an artery.

It is envisaged by the present invention that the various recombinant DNA molecules and vectors of the invention are administered either alone or in any combination using standard vectors and/or gene delivery systems, and optionally together with an appropriate compound, for example VEGF, and/or together with a pharmaceutically acceptable carrier or excipient. Subsequent to administration, said recombinant DNA molecules may be stably integrated into the genome of the mammal. On the other hand, viral vectors may be used which are specific for certain cells or tissues, preferably for the endothelium and persist in said cells. Suitable pharmaceutical carriers and excipients are well known in the art. The pharmaceutical compositions prepared according to the invention can be used for the prevention or treatment or delaying of different kinds of diseases, which are related to the expression or overexpression of a given gene or genes in the endothelium.

Furthermore, it is possible to use a pharmaceutical composition of the invention which comprises a recombinant DNA molecule or vector of the invention in gene therapy. Suitable gene delivery systems may include liposomes, receptor-mediated delivery systems, naked DNA, and viral vectors such as herpes viruses, retroviruses, adenoviruses, and adeno-associated viruses, among others. Delivery of nucleic acids to a specific site in the body for gene therapy or antisense therapy may also be accomplished using a biolistic delivery system, such as that described by Williams (Proc. Natl. Acad. Sci. USA 88 (1991), 2726-2729).
Standard methods for transfecting cells with recombinant DNA are well known to those skilled in the art of molecular biology, see, e.g., WO 94/29469. Gene therapy

and antisense therapy to prevent or decrease the development of atherosclerosis or inhibit angiogenesis may be carried out by directly administering the recombinant DNA molecule or vector of the invention to a patient or by transfecting endothelial cells with the recombinant DNA molecule or vector of the invention *ex vivo* and infusing the transfected cells into the patient. Furthermore, research pertaining to gene transfer into cells of the germ line is one of the fastest growing fields in reproductive biology. Gene therapy, which is based on introducing therapeutic genes into cells by ex-vivo or in-vivo techniques is one of the most important applications of gene transfer. Suitable vectors and methods for in-vitro or in-vivo gene therapy are described in the literature and are known to the person skilled in the art; see, e.g., WO94/29469, WO 97/00957 or Schaper (Current Opinion in Biotechnology 7 (1996), 635-640) and references cited therein. The DNA molecules and vectors comprised in the pharmaceutical composition of the invention may be designed for direct introduction or for introduction via liposomes, or viral vectors (e.g. adenoviral, retroviral) containing said recombinant DNA molecule into the cell. Preferably, said cell is a germ line cell, embryonic cell, or egg cell or derived therefrom. The pharmaceutical compositions according to the invention can be used for the treatment of kinds of diseases hitherto unknown as being related to the expression and/or over expression of genes in the endothelium.

The present invention also relates to diagnostic compositions or kits comprising at least one of the aforementioned recombinant DNA molecules or vectors, and in the case of diagnostic compositions, optionally suitable means for detection.
Said diagnostic compositions may be used for methods of detecting and isolating regulatory sequences which are a functionally equivalent to the Flk-1 intron regulatory sequences of the invention. The kits of the invention may further contain compounds such as further plasmids, antibiotics and the like for screening transgenic animals and/or animal cells useful for the genetic engineering of non-human animals, preferably mammals and most preferably mouse.

It is to be understood that the introduced recombinant DNA molecules and vectors of the invention express the heterologous DNA sequence after introduction into said cell and preferably remain in this status during the lifetime of said cell. For example, cell lines which stably express the heterologous DNA under the control of the regulatory sequence of the invention may be engineered. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with the recombinant DNA molecule or vector of the invention and a selectable marker, either on the same or separate vectors. Following the introduction of foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines which express the heterologous DNA sequence under the control of the regulatory sequence of the invention, and which respond to VEGF and/or hypoxia mediated signal transduction. Such engineered cell lines are particularly useful in screening compounds capable of modulating Flk-1 gene expression.

A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler, Cell 11(1977), 223), hypoxanthine-guanine phosphoribosyltransferase (Szybalska, Proc. Natl. Acad. Sci. USA 48 (1962), 2026), and adenine phosphoribosyltransferase (Lowy, Cell 22 (1980), 817) genes can be employed in tk, hgprt or aprt cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for dhfr, which confers resistance to methotrexate (Wigler, Proc. Natl. Acad. Sci. USA 77 (1980), 3567; O'Hare, Proc. Natl. Acad. Sci. USA 78 (1981), 1527), gpt, which confers resistance to mycophenolic acid (Mulligan, Proc. Natl. Acad. Sci. USA 78 (1981), 2072); neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin, J. Mol. Biol. 150 (1981), 1); and hygro, which confers resistance to hygromycin (Santerre, Gene 30 (1984), 147) genes. Additional selectable genes have been described, namely trpB, which allows cells to utilize indole in place of tryptophan; hisD, which allows cells to utilize histinol in place of histidine (Hartman, Proc. Natl. Acad. Sci. USA 85 (1988), 8047); and ODC

(ornithine decarboxylase) which confers resistance to the ornithine decarboxylase inhibitor, 2-(difluoromethyl)-DL-ornithine, DFMO (McConlogue, 1987, In: Current Communications in Molecular Biology, Cold Spring Harbor Laboratory ed.). On the other hand, the person skilled in the art may also use the regulatory sequences of the invention to "knock out" an endogenous gene comprising identical or similar regulatory sequences, for example, by gene targeting, cosuppression, triple helix, antisense or ribozyme technology.

The present invention also relates to a method for the production of a transgenic animal, preferably mouse, comprising introduction of a recombinant DNA molecule or vector of the invention into a germ cell, an embryonic cell or an egg or a cell derived therefrom. The non-human animal to be used in the method of the invention may be a wildtype, i.e. healthy animal, or may have a disease or disorder, preferably a disease or disorder which is dependent on neovascularization, such as solid tumors, retinopathy, arthritis, psoriasis. Said disease or disorder may be an inborne insufficiency or naturall developed or caused by genetical engineering, for instance by the expression of a DNA sequence encoding a protein involved in neuronal development and/or diseases as described above, preferably under the control of the regulatory sequence of the invention.

The invention also relates to transgenic non-human animals comprising a recombinant DNA molecule or vector of the invention or obtained by the method described above, preferably wherein said recombinant DNA molecule is stably integrated into the genome of said non-human animal, preferably such that the presence of said recombinant DNA molecule or vector leads to the transcription and/or expression of the heterologous DNA sequence by the regulatory sequence of the invention. Further non-human animals which may be employed according to the embodiments of the invention as described above are well known to the person skilled in the art and comprise rat, hamster, dog, monkey, rabbit, pig.

With the regulatory sequences of the invention, it is now possible to study in vivo the regulation of Flk-1 expression during angiogenesis. Furthermore, since VEGF and VEGF receptor genes seem to have different functions in different stages of development, it is now possible to determine domains of said proteins which may be important for their biological activity and/or for the regulation of their activity. In addition, it is now possible to in vivo study mutations which affect different functional or regulatory aspects of VEGF or its receptor or vector of the invention.

Moreover, the present invention relates to a method for the identification of a chemical and/or biological substance capable of suppressing or activating and/or enhancing the transcription of a gene in endothelial cells comprising:

(a) contacting a cell of the invention or the transgenic non-human animal of the invention either of which is capable of expressing the heterologous DNA sequence with a plurality of compounds; and
(b) determining those compounds which suppress or activate and/or enhance the expression of said heterologous DNA sequence.

Said plurality of compounds may be comprised in, for example, samples, e.g. cell extracts from, e.g. plants, animals or microorganisms. Furthermore, said compounds may be known in the art but hitherto not known to be capable of suppressing or activating and/or enhancing the transcription of a gene in endothelial cells. The plurality of compounds may be, e.g., added to the culture medium or injected into the animals.

The term "plurality of compounds" in a method of the invention is to be understood as a plurality of substances which are either identical or not. If a sample containing a plurality of compounds is identified in the method of the invention, then it is either possible to isolate the compound from the original sample identified as containing the compound capable of suppressing or activating and/or enhancing the transcription of a gene in endothelial cells, or one can further subdivide the original sample, for example, if it consists of a plurality of different compounds, so as to reduce the number of different substances per sample and repeat the method with the subdivisions of the original sample. Depending on the complexity of the samples,

this can be done several times, preferably until the sample identified according to the method of the invention only comprises a limited number of or only one substance(s). Preferably said sample comprises substances of similar chemical and/or physical properties, most preferably said substances are identical.

Determining whether a compound is capable of suppressing or activating and/or enhancing the transcription of a gene in endothelial cells can be done, for example, in mice by monitoring reporter gene expression or by monitoring behavior of the transgenic non-human animals of the invention contacted with the compounds compared to that of wild-type animals or compared to a transgenic non-human animal contacted with a compound which is either known to be capable or incapable of suppressing or activating and/or enhancing the transcription of a gene in endothelial cells of said transgenic non-human animal of the invention. Furthermore, the person skilled in the art can monitor the physical behavior, or for example the movement of the above-described animals. Such methods are well known in the art.

Such regulators of Flk-1 gene expression may be used in processes such as wound healing; in contrast, antagonists of expression may be used in the treatment of tumors that rely on vascularization for growth. Thus, the present invention provides methods for identifying compounds which modulate VEGF receptor (e.g., Flk-1 or Flt1) gene expression. Compounds found to downregulate expression of a VEGF receptor gene can be used in methods to inhibit angiogenesis, while compounds found to enhance Flk-1 or Flt1 expression can be used in methods to promote angiogenesis, for example, to promote wound healing (e.g., healing of broken bones, burns, diabetic ulcers, and traumatic or surgical wounds) or to treat peripheral vascular disease, atherosclerosis, cerebral vascular disease, hypoxic tissue damage (e.g., retinopathy, hypoxic damage to heart tissue), diabetic pathologies such as chronic skin lesions, or coronary vascular disease. These compounds can also be used to treat patients who have, or have had, transient ischemic attacks, vascular graft surgery, balloon angioplasty, frostbite, gangrene, or poor circulation. Compounds identified as having the desired effect (i.e., enhancing or inhibiting Flk-1 expression) can be tested further in appropriate models of endothelial cell growth and angiogenesis which are known to those skilled in the art.

The therapeutic compounds identified using the method of the invention may be administered to a patient by any appropriate method for the particular compound, e.g., orally, intravenously, parenterally, transdermally, transmucosally, or by surgery or implantation (e.g., with the compound being in the form of a solid or semi-solid biologically compatible and resorbable matrix) at or near the site where the effect of the compound is desired. For example, a salve or transdermal patch that can be directly applied to the skin so that a sufficient quantity of the compound is absorbed to increase vascularization locally may be used. This method would apply most generally to wounds on the skin. Salves containing the compound can be applied topically to induce new blood vessel formation locally, thereby improving oxygenation of the area and hastening wound healing. Therapeutic doses are determined to be appropriate by one skilled in the art.

Furthermore, identification of transacting factors which interact with the regulatory sequences of the invention can form the basis for the development of novel therapeutics for modulating conditions associated with endothelial cell growth, such as angiogenesis, vascular disease, and wound healing. Identification of transacting factors is carried out using standard methods in the art (see, e.g., Sambrook, supra, and Ausubel, supra). To determine whether a protein binds to the regulatory sequences of the invention standard DNA footprinting and/or native gel-shift analyses can be carried out. In order to identify the transacting factor which binds to the regulatory sequence of the invention, the regulatory sequences can be used as an affinity reagent in standard protein purification methods, or as a probe for screening an expression library. Once the transacting factor is identified, modulation of its binding to the regulatory sequence in the Flk-1 gene can be pursued, beginning with, for example, screening for inhibitors of transacting factor binding. Enhancement of Flk-1 expression in a patient, and thus enhancement of angiogenesis, may be achieved by administration of the transacting factor, or the gene encoding it (e.g., in a vector for gene therapy). In addition, if the active form of the transacting factor is a dimer, dominant-negative mutants of the transacting factor could be made in order to inhibit its activity. Furthermore, upon identification of the

transacting factor, further components in the pathway of Flk-1 signal transduction can be identified. Modulation of the activities of these components can then be pursued, in order to develop additional drugs and methods for modulating endothelial cell growth and angiogenesis.

As discussed in the background section of the description of the present invention, the interaction of VEGF and its receptor play an important role in the onset of angiogenic disease. Transgenic non-human animals expressing VEGF and/or its receptor gene and/or mutated versions thereof under the control of the regulatory sequences of the invention can now be used for the identification of substances, which, for example, are capable of restoring the wild-type interaction of mutated VEGF and its receptor either or both of which bear mutations. Some genetic changes lead to altered protein conformational states. Genetic changes may therefore result in a decreased binding activity of VEGF. Restoring the activity of mutant VEGF protein or increasing the activity of other proteins which interact with mutant VEGF proteins is the most elegant and specific means to correct these molecular defects. In addition, some genetic changes may result in altered conformational states of the receptor. This, in turn, may functionally inactivate the tyrosine kinase activity, making it incapable of signal transduction. In order to restore the function of such mutant proteins an antibody may be used which binds to an epitope and induces a conformational change of the protein thereby restoring the wild type function. Thus, the methods of the invention are also useful to screen e.g., antibody, Fab, Fv or scFv expression libraries wherein the DNA sequence encoding said antibodies or derivatives thereof are under the control of the regulatory sequence of the invention. It is, of course, evident to the person skilled in the art that also other protein or peptide expression libraries using the regulatory sequences of the invention may be employed.

Further, the present invention relates to the use of the recombinant DNA molecule, vector, cell, pharmaceutical compositions, diagnostic compositions or a transgenic non-human animal of the invention for the identification of a chemical and/or

biological substance capable of suppressing or activating and/or enhancing the transcription, expression and/or activity of genes and/or its expression products in endothelial cells.

In a preferred embodiment, the chemical or biological substance used in the methods and uses of the present invention is selected from the group consisting of peptides, proteins, nucleic acids, antibodies, small organic compounds, hormones, neural transmitters, peptidomimics, and PNAs (Milner, Nature Medicine 1 (1995), 879-880; Hupp, Cell 83 (1995), 237-245; Gibbs, Cell 79 (1994), 193-198).

The present invention further relates to a method of inhibiting a vascular disease in a subject, comprising contacting an artery of said subject with the recombinant DNA molecule or vector of the invention, wherein said protein reduces or prevents the development of the vascular disease, preferably said protein reduces proliferation of smooth muscle cells.

In a further embodiment the present invention relates to the use of a recombinant DNA molecule, vector, nucleic acid molecule of the invention and/or substance identified by a method of the invention for the preparation of a composition for directing and/or preventing expression of genes specifically in endothelial cells and/or for the preparation of a pharmaceutical composition for treating, preventing and/or delaying a vascular disease and/or a tumorous disease in a subject.

In a further embodiment, the present invention relates to the use of a recombinant DNA molecule, vector and/or the nucleic acid molecule of the invention for the preparation of a pharmaceutical composition for inducing a vascular disease in a non-human animal or in a transgenic non-human animal described above.

In a preferred embodiment of the methods and uses of the invention, the vascular disease is atherosclerosis and/or a neuronal disorder. Further possible methods and uses in accordance with the present invention will be evident to the person skilled in

the art and are described in, for example, WO 95/13387, WO 94/11499 and WO 97/00957.

The recombinant DNA molecules, vectors, nucleic acid molecules, compounds, uses and methods of the invention can be used for the treatment of all kinds of disorders and diseases hitherto unknown as being related to or dependent on the modulation of genes specifically expressed in the endothelium. The recombinant DNA molecules, vectors, nucleic acid molecules, compounds, methods and uses of the present invention may be desirably employed in humans, although animal treatment is also encompassed by the methods and uses described herein. Thus, the present invention provides for the use of a regulatory sequence as defined above for enhancing and/or directing gene expression in endothelial cells in any kind of organism.

The figures show:

Figure 1: Nucleotide sequence of the murine Flk-1 gene. The ATG codon is at position +299. The three exons are indicated in bold. Motifs for transcription factors are underlined. VRE: vascular response element,

Figure 2: Map of reporter gene construct pGL2-B. Arrows symbolize functional elements. Luc: luciferase gene, AMP: ampicillin resistance gene, f1ori: replication origin for bacteriophage f1.

Figure 3: Map of reporter gene construct pGLacZ. Arrows symbolize functional elements. LacZ: β-galactosidase gene, AMP: ampicillin resistance gene, f1ori: replication origin or bacteriophage f1.

Figure 4: Functional analysis of the Flk-1 intron. A) Schematic representation of the first 6.5 kbp of the transcribed Flk-1 gene. The first three exons are depicted as hatched boxes. Bottom: Subdivision of the region into

three fragments (5'-In1, 3'-In1, In2). B) Transfection assay of the intron fragments in combination with the Flk-1 promoter region of bp -640 to bp +299. The values were coordinated with 5'-In1 fragments with respect to the activity of the construct. C) Transfection assay of the intron fragments in combination with the Flk-1 promoter region between nucleotides -4.1 kbp to +299 bp.

Figure 5: Analysis of the intron in transgenic mice. The embryo (10.5 days) was stained overnight with X-Gal. The reporter gene was under the control of the intron enhancer (3'-In1, cf. Fig. 4A) and of the Flk-1 promoter fragment ranging from nucleotides -4100 to +299. A) Top lateral view. B) Top dorso-cranial view.

Figure 6: Histological evaluation of a transgenic embryo. The embryo shown in Fig. 5 was embedded in paraffin. The cuts were stained with neutral red. A) Pseudo transversal cut through the head region. B) Magnification of a section from A. C) Pseudo transversal cut of a caudal region. 1: 4th ventricle of cerebrum, 2: acoustic vesicle/otocyte, 3: V. cardinalis anterior. 4: third ventricle of cerebrum. 5: endbrain vesicle. 6: optic vesicle. 7: ganglion trigeminale (V). 8: chorda dorsalis.

Figure 7: Functional analysis of the first two introns of the Flk-1 gene in vivo. The depicted embryo (10.5 days) carries the ß-galactosidase gene under the control of the Flk-1 promoter (-4.1 kbp to +299 bp) and the first 6.2 kbp of the transcribed region (cf. Fig. 4A). The staining was carried out as described in Fig. 5.

Figure 8: In vivo characterization of the intron enhancer in combination with the strongest promoter fragment. All three embryos carry the ß-galactosidase gene under the control of the Flk-1 promoter fragment

of bp -640 to bp +299 and the intron enhancer. The staining was carried out as described in Fig. 5.

Figure 9: Detailed analysis of the left-hand embryo from Fig. 8A) Left lateral view. B) Sectional magnification of A. C) Right lateral view.

Figure 10: Histological evaluation of the embryo depicted in Fig. 9. The embryo was embedded in paraffin and was cut into 10 μm slices. The cuts were stained with neutral red. A) Pseudo transversal cut through the head region. B) Magnification from a similar cut level as in A. C) Pseudo transversal cut from a more caudally located section. D) Pseudo transversal cut from thoracal section. 1: 4th ventricle cerebrum, 2: 3rd ventricle cerebrum, 3: endbrain vesicle, 4: A. carotis interna, 5: ganglion trigeminale (V), 6: V. cardinalis anterior, 7: neural tube, 8: esophagus, 9: V. cardinalis posterior, 10: aorta dorsalis, 11: endocardium of the heart atrium, 12: vessels of the myocardium.

The examples illustrate the invention.

Example 1: Cloning and construction of Flk-1 intron/reporter gene vectors

DNA clones containing the 5' region of the mouse Flk-1 gene were isolated from a library prepared from 129/SvJ mice in λ Dash II vector (Stratagene) (Rönicke, supra) or in λ FIX II or obtained from the P1 Library (Genome Systems, St. Louis). A 21 kb region of the mouse Flk-1 gene, contained in the DNA insertions of two λ phages 6 and 16, including approximately 15 kb of 5' flanking sequences, exons 1, 2 and 3, and introns 1 and 2 was characterized by restriction enzyme mapping and Southern blot analysis. Lower DNA fragments of the phage clones were cloned into pBluescript vector DNA (Stratagene) and used for further characterization. Sequencing was performed using an automatic Sequencer (373A, Applied Biosystems). The DNA sequence (SEQ ID NO: 1) of a 12.8 kb region spanning from

position -6,660 kb (relative to the transcriptional start site) to position +6,135 kb (located in the third exon) was determined (Fig. 1). Figure 4A shows a schematic representation of the first 6.5 kbp of the transcribed region of the murine Flk-1 gene. Exons I, II and III are emphasized as hatched boxes. The first intron having a length of 3.6 kbp is subdivided into two regions (5'-In1 and 3'-In1). The region In-2 contains the entire second intron, the second exon, the 3' end of the first intron and part of the third exon. This subdivision into various intron fragments was maintained in the following analyses. The reporter gene constructs used were derived from pGL2 basic vector (Promega) that contains a promoterless luciferase gene. Luciferase reporter gene constructs were generated for transfection of cells in vitro. For use in transgenic mice in vivo, plasmids were used in which the luciferase reporter gene was replaced by a lacZ reporter gene.

In order to generate (luciferase) reporter gene constructs, Flk-1 promoter fragments were amplified by PCR and cloned into pGL2 (Promega) vector DNA 5' to the luciferase gene as described by Rönicke, supra; see also Figure 2. In short, the upstream primers used were -1900: 5'-GGG GTA CCG AAT TCT AAA TGG GGC GAT TAC C-3' (SEQ ID NO 2); -640: 5'-GTG GTA CCC AAA CAC TCA ACA CCA CTG-3'(SEQ ID NO: 3); -624, 5'-TCG GTA CCG ACC CAG CCA GGA AGT TC-3' (SEQ ID NO: 4); the downstream primer was +299, 5'-TTG CTA AGC TTC CTG CAC CTC GCG CTG GG-3' (SEQ ID NO: 5). To generate the construct ranging from -4100 to +299, a HindIII-EcoRI fragment of recombinant lambda phage 6 from P1 Library (Genome Systems, St. Louis) was inserted into the plasmid ranging from -1900 to +299. Vectors that contained Flk-1 intron sequences in addition to promoter sequences were generated as follows: specific intron sequences were amplified by PCR from cloned Flk-1 genomic DNA and inserted downstream of the reporter gene. Primers used for amplification were 5'-In1down: 5'-AGG GAT CCA CTC TTT AGT AGT AAG GCG-3' (nucleotides 7036-7057 of SEQ ID NO: 1, SEQ ID NO: 6); 5'-In1up: 5'-ACC TCG AGA CTT GGA TGG CAC-3' (nucleotides 8324-8342 of SEQ ID NO: 1, SEQ ID NO: 7); 3'-In1down: 5'-GGG CTA TAA TTG GTG CCA TCC-3' (nucleotides 8312-8332 of SEQ ID NO: 1, SEQ ID NO: 8); 3'-In1up: 5'-GGA TGG AGA AAA TCG CCA GGC-3' (nucleotides 10637-10658 of SEQ ID NO: 1, SEQ ID

NO: 9); IN2A: 5'-GTG TGC ATT GTT TAT GGA AGG G-3' (nucleotides 10571-10593 of SEQ ID NO: 1, SEQ ID NO: 10); IN2B: 5'-CAT AGA CAT AAA CAG TGG AGG C-3' (nucleotides 12849-12871 which is part of the cDNA sequence published by Millauer (1993), supra, SEQ ID NO: 11). For the subsequent experiments the vector indicated in Fig. 3 was used. It represents a modification of the pGL2 basic vector in which the corresponding Flk-1 promoter fragments were inserted into the KpnI and HindIII restriction sites of the polylinker (Fig. 2). Also the luciferase reporter gene was replaced by the ß-galactosidase gene (Schlaeger, Proc. Natl. Acad. Sci. USA 94 (1997), 3058-3063). For an analysis of the intron intron fragments were cloned into the BamHI and SalI restriction sites indicated. DNA manipulations, PCR amplification and DNA sequencing were performed according to conventional methods known in the art as described, for example in Sambrook, supra and PCR Technology, Griffin and Griffin, eds., RC Press London (1994).

Example 2: Functional analysis of the intron of the Flk-1 gene in vitro

Figure 4 shows the result of transient transfections in BAECs. The corresponding intron fragments were combined with a Flk-1 promoter fragments which comprised nucleotides -640 to +299. The promoter activity was standardized with respect to the promoter activity of the construct containing the 5'-In1 fragment.

Tissue culture and transient transfections were performed as follows:

All cells were cultured in DMEM+ supplemented with 10% FCS (Sigma). bEnd5 cells were generated by transformation with the Polyoma middle-T oncogene as described earlier (Montesano, Cell 62 (1990), 435-445). BAECs were prepared as described (Schwartz, In Vitro 14 (1978), 966-980). NIH 3T3, C2C12 and L cells were obtained from ATCC. Transient transfections were performed using the CaPO4-precipitation method according to Chen and Okayama (Mol. Cell Biol. 7 (1987), 2745-2752) . Each construct was transfected at least six times in three independent experiments. Cells were grown to 70% confluence in 6-cm dishes prior to transfection. Cells were washed 16 hrs after addition of CaPO-precipitate and incubated for further 48 hrs. In each experiment, 6 μg luciferase and 1 μg pCMV5

(Rönicke, supra) lacZ reporter gene constructs were used. Cells were lysed in 1 x reporter-lysis-buffer (Promega) for 15 min on a test tube-rotator. After centrifugation, the supernatant was transferred to a fresh tube and stored at -80 °C or taken for luciferase-and lacZ-assay immediately. Reporter-gene assays for β-galactosidase activity were performed according to Eustice (Biotechniques 11 (1991), 739-740). Chlorophenol red-β-D-galactopyranoside (CPRG) was used as a substrate and the conversion was measured at 575 nm in an ELISA-reader (Biometra). Extracts were diluted to obtain OD575nm values between 0.2 and 0.8. These values were used to standardize for transfection efficiency after subtracting a background value, as determined from a cell extract of a transfection without lacZ-reporter plasmid but with a luciferase-reporter plasmid. Luciferase-reporter gene assays were performed with the same extracts as described by the manufacturer (Promega). Luciferase activity was measured with a luminometer (LB96P, Berthold) and calculated as per cent of the activity of the pGL2-promoter plasmid (Promega).

Construct	5'-In1	3'-In1	In2
BAEC	100+/-0%	128+/-34%	136+/-52%
3T3	100+/-0%	55+/-15%	74+/-33%

Table I: Functional analysis of the intron of the Flk-1 gene. The upper line indicates the corresponding intron fragment which was analyzed in combination with the Flk-1 promoter (-640 bp/+299 bp).

Figure 4C shows the results of another transfection assay of the intron fragments. It was carried out as described above, with the exception that a Flk-1 promoter fragments was used that comprised the region between nucleotides -4100 and +299. Also, a fragment was analyzed that contained the entire first intron, the second exon, the second intron and part of the third exon shown in Figure 4A.

Construct	5'-In1	3'-In1	In2	In1+2
BAEC	100+/-0%	206+/-81%	119+/-51%	154+/-68%
3T3	100+/-0%	71+/-32%	85+/-27%	35+/-12%

Table II.: Functional analysis of introns of the Flk-1 gene. The upper line indicates the corresponding intron fragment which was analyzed in combination with the Flk-1 promoter (-4100 Bp/+299 Bp).

An analysis of this experiment revealed that the construct with the 3' region of the first intron in BAECs had an activity that was twice that of that containing the 5' region of the first intron. Also, it showed 85% higher activity than the construct with the second intron (p=0.0153). The 4.5 kbp longer construct In1+2 that also contained the 3' region of the first intron, too, revealed an activity that was markedly higher in BAECs than in 3T3 cells.

A functional analysis of the first 6.5 kbp of the transcribed region of the murine Flk-1 genes lead to the identification of an endothelial-specific positive regulatory element. This regulatory sequence is located in the region between the XhoI and BamHI restriction site in the first intron of the Flk-1 gene (cf. Fig. 4A). It is functional in both orientations since the intron is used in an antiparallel manner with respect to the Flk-1 promoter fragment in the construct referred to as 3'-In 1. In construct In1+2, however, the original orientation was maintained. A sequence analysis of the intron enhancer lead to the identification of two potential GATA binding sites at position +1927 bp and +3514 bp; (Evans, Proc. Natl. Acad. Sci. USA 85 (1988), 5976-5980; Orkin, Blood 80 (1992), 575-581), a potential AP-1 binding site at position +2210 bp; (Lee, Cell 49 (1987), 741-752) and two PEA3 consensus sequences at position +3494 bp and +3741 bp; (Martin, Proc. Natl. Acad. Sci. USA 85 (1988), 5839-5843).

Example 3: Functional characterization of the Flk-1 promoter in vivo.

So far, analyses of the murine Flk-1 promoter have been restricted to in vitro systems (Rönicke, supra; Patterson, supra). The investigation of the promoter activity in vitro is an important tool in promoter characterizing since it is useful to assay a large number of promoter constructs for their activity in a short time. However, this situation is always an artificial one since not all factors that are relevant in vivo can also be reconstituted in vitro. While an in vitro investigation of a promoter yields important information on the mechanisms of gene regulation it is only the in vivo characterization that can yield the final proof for the relevance of the elements identified. An excellent test system for promoter analysis in vivo are transgenic mice. In this model the corresponding promoter fragment was cloned before a reporter gene, isolated together with this reporter gene and injected into fertilized mouse oocytes. In many cases, successful integration of the promoter reporter construct into the mouse genome lead to transgenic mice which contain the construct in every cell. This test system, in addition to the analysis of the promoter activity during embryonic development and in the adult animal, allows a tissue-specific characterization of the promoter activity.

For the investigation of the Flk-1 promoter in transgenic mice the bacterial ß-galactosidase reporter gene was chosen since the gene product is easily detectable by color reaction and remains at the location of production due to its limited solubility. In this manner it is possible to identify cells in which the corresponding Flk-1 promoter fragment since only there an expression of ß-galactosidase took place.

When producing transgenic mice it was taken care that no regions originating from the vector were injected along with the promoter. First, Flk-1 promoter fragments comprising the regions between nucleotides -640 and +299, -1900 and +299 as well as -4100 and +299 were investigated. The constructs were based on plasmid pGL-2B described in Figure 2 with the exception that the luciferase reporter gene was replaced by the ß-galactosidase gene. All injection fragments used in the examples were obtained by restriction digestion with the enzymes KpnI and SalI. Transgenic mice were generated as described by (Hogan, Manipulating the Mouse Embryo

(1994), Cold Spring Harbor Laboratory Press, New York). The embryos to be examined were isolated on day 10 after reimplantation of the injected oocytes. Analysis of the transgenic embryos revealed that although promoter activity could be detected, none of the constructs was capable of conferring reproducible expression of the reporter gene in the endothelium.

Example 4: Functional characterization of the Flk-1 intron in vivo

After analysis of the Flk-1 promoter region from -4.1 kbp to +299 Bp the intron which was identified in vitro was then examined for its function in vivo. For this purpose, a construct similar to that shown in Figure 3 was used which contained an Flk-1 promoter fragment ranging from nucleotide -4100 to base pair +299 and the intron enhancer (3'-In1, cf. Fig. 4A). The staining and fixation of the embryos was performed as follows: The mid-day of the plug observation was counted as E0.5. The embryos were dissected out in ice-cold PBS and fixed in ice-cold 2 % paraformaledehyd, 2 mM $MgCl_2$, 2 mM EGTA, 0.1 M Pipes buffer, pH 6.9 for 15 minutes. The embryos were rinsed with PBS three times for 5 minutes each. The LacZ expression was detected by incubating the embryos at 30 °C overnight in 0.1 % X-gal, 5 mM potassium ferricyanide, 5 mM potassium ferrocyanide, 1 mM magnesium chloride, 0.002 % NP-40, 0.01 % sodium deoxycholate, PBS, pH 7.0. After the staining, embryos were rinsed in PBS and postfixed at 4 °C overnight in 2 % paraformaldehyde, 0.1 % glutaraldehyde, PBS, pH 7.0. For whole-mount photography, the postfixed embryos were rinsed in PBS and equilibrated in 50 % glycerol and then in 70 % glycerol. Figure 5 shows an embryo (embryonic day 10.5) which was isolated after injection of the fragment. In Figure 5A a color reaction in vessels of the developing brain can be clearly discerned. Also superficial vessels in the body's middle (dorsal) and a staining in the liver bud can be observed. Figure 5B shows the dorsal, caudal region of the same embryo. It proves that the vessels in both halves of the head were stained.

For an exact localization of the stained cells the embryo was embedded in paraffin, cut into slices and counterstained with neutral red. The results of the histological

analysis are shown in Figure 6. It shows pseudo transversal cuts of the embryo. In section 6A a staining of the inner lining of the V. cardinalis anterior (3) and of other superficial vessels can be seen. Figure 6B represents a strong magnification of a section of 6A with the staining of endothelial cells within the V. cardinalis anterior. Figure 6C shows a more caudally located region. Again, the staining of the V. cardinalis anterior and of superficial vessels with wide lumen and thin walls as well as of vascular structures in the neural tube is clearly visible. Also, a staining of the chorda dorsalis (9) can be observed. However, in none of the cases a staining of arterial vessels could be observed.

The subsequent injection of the same fragment lead to a total of eight further transgenic embryos which displayed an identical expression pattern albeit in two cases of weaker nature. Thus, the intron enhancer exhibited in vivo an effect that was even more marked than in vitro. In combination with a promoter fragment which on its own had a very variable expression pattern it ensures a reproducible expression pattern with clear endothelium specificity, however, covering a substantial part of the endogenous Flk-1 expression pattern.

Example 5: Functional analysis of the introns I and II of the murine Flk-1 gene in vivo

Since the intron enhancer in combination with an Flk-1 promoter fragment displayed an endothelium-specific function in transgenic mice covering a substantial part of the endogenous expression pattern, the further search for in vivo relevant, gene regulatory elements was extended to other intron regions. For this purpose, the construct containing the promoter region between nucleotides -4100 to +299 and the first 6.5 kbp of the transcribed region (In1+2; cf. Fig. 4) was used. Figure 7 shows an embryo on embryonic day 10.5 which was obtained after injection of this fragment. Again, a staining of the vessels in the developing brain as well as superficial vessels the of the liver bud was visible. The following injections yielded four further transgenic embryos which displayed the same pattern. A combination of the

promoter region used with only the 5' end of the first intron (5'-In1; cf. Fig. 4), however, yielded no endothelium-specific expression pattern.

Example 6: Combination of the intron enhancer with the Flk-1 promoter fragment that was the most potent in vitro

To investigate whether the repressing elements of the murine Flk-1 promoter between nucleotides -4100 and -640 are functional also in combination with the intron enhancer, a shorter construct without these inhibitory regions was used for further analysis. It contain the intron enhancer (3'-In1) and the 5' region from base pair -640 to nucleotide +299. This 5' region displayed the highest activity in vitro. Figure 8 shows three transgenic embryos (embryonic day 10.5) which were obtained after injection of the fragment. All three display a more marked staining in vascular structures than the embryos analyzed so far. While the embryo on the right hand shows a weak staining, the left-hand embryo yields a very strong expression in virtually all vessels. The embryo in the middle holds a medium position as regards the completeness of its expression pattern, i.e., it lacks expression in the heart although it resembles strongly the embryo on the left hand as regards the staining of the other structures. In Figure 9A the left-hand embryo from Fig. 8 is shown in more detail. The strong staining of the heart in the region of the atrium and ventricle is particularly clearly visible. Furthermore, the vessels of the developing brain, the vessels between the somites, the aorta dorsalis as well as the fine capillary plexus on the body's surface are stained. Figure 9B shows a sectional magnification of 9A. Here, the staining of the vessels in the head region as well as the expression in the superficial capillary plexus is visible. In Figure 9C the same embryo is shown from the other side. In addition to the structures described in Figure 9A also a staining of the chorda dorsalis can be observed.

The embryo shown in Figure 9 was embedded in paraffin and cut to slices. The cuts dyed with neutral red are shown in Fig. 10. Figure 10A shows a pseudo transversal cut through the head region. Particularly prominent is the branching of the A. carotis interna (4) in addition to the staining of other vascular structures. Figure 10B

represents a magnification of a similar cut; here, too, the branching of the A. carotis interna is particularly striking. Figure 10C shows a more caudal cut which in terms of its position roughly corresponds to the cut shown in Figure 6A. Here, however, in addition to the staining of the V. cardinalis anterior (6) an expression in the branching of the A. carotis interna (4) and other vascular structures is visible. Figure 10D represents an even more caudally located region. A staining in the venous endothelium (V. cardinalis posterior, 9) and in the arterial structures (aorta dorsalis, 10) can be observed. Furthermore, the endocardium of the atrium as well as the vessels in the trabeculae of the heart ventricles display an expression.

A total of seven transgenic embryos was analyzed after injection of this fragment (-640 bp/+299 bp/3'-In1). Safe for one which showed no staining, all embryos displayed an expression of the ß-galactosidase in endothelial structures. The staining was regularly more marked than in combination with the negative regulatory elements between nucleotides -4100 and -640. Thus, the in vitro identified regions displayed a function in vivo. The deletion of these negative regulatory elements yielded a construct that lead to a reproducible expression in venous and arterial endothelium.

In summary, reporter gene constructs that comprise a Flk-1 intron fragment (+1570 to +3997) downstream of the lacZ gene show reproducibly reporter gene expression in the vasculature of transgenic mice.

The present invention is not to be limited in scope by its specific embodiments described which are intended as single illustrations of individual aspects of the invention and any DNA molecules, or vectors which are functionally equivalent are within the scope of the invention. Indeed, various modifications of the invention in addition to those shown and described therein will become apparent to those skilled in the art from the foregoing description and accompanying drawings. Said modifications intended to fall within the scope of the appended claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Max-Planck-Gesellschaft zur Foerderung der Wissenschaften e. V.
(B) STREET: none
(C) CITY: Berlin
(D) STATE: none
(E) COUNTRY: Germany
(F) POSTAL CODE (ZIP): none

(ii) TITLE OF INVENTION: Regulatory Sequences Capable Of Conferring Expression Of A Heterologous DNA Sequence In Endothelial Cells In Vivo And Uses Thereof

(iii) NUMBER OF SEQUENCES: 11

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12845 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
TCTAGAATAT AGAAGATAAG TTTGCGTACA ATTCAGTCCT TTGAAGACCT GATAAGCTTT     60
AAGAAGGAAG ATGGGTTACA CATTGGGAAA TGGTTGCAAT CTGCACATGG CAGAGGCAAG    120
AGATGCAAAT CACATTTCTT ACATACTCCA TACAAATCTT ACAAGACTGT TTTTCTTTCT    180
CATTTAAAAT AAGAAGACCT GCCAGTCTTC CCCTTATTAC TAATTACAGT CACTCTGTAT    240
CTTTGTTGAC ATTGGATAGT TTTACATACT TCAACAGGCT GGTGTCATTA AAGTTGTGGT    300
GGGTGGGCAC CAGAGACACG TGATTCAGAG TGGGAGGAGA TGCAGGAGAA ACGAGGCACA    360
```

GCAGAAGCAG AAGCGAGGAA AAACACTCTC AACGTTACTA ACACATCGAG AGGTTCCGCA 420

CACTAGCAAT ACGGGCTGAA TCTGACCTAA TCTCTGCTGT TGAAAATTTT GCCTAGCCGC 480

ACACTAGCAA TACGGGCTGA ATCTGACCTA ATCTCTGCTG TTGAAAATTT TGCCTAGCCT 540

GTCACACAAG TGCTGAGCAT ACAGAAAAAG GAGAGTAATT CTCTGGTTCT TTGACTAACC 600

AAATAGTCTA TATCAAATTG CCTAAGATAA TGTATACATT TAGTACATGA CTGGTTATAC 660

CTATTCTATA TGACTATTAT TTAAATGTGA ATTTACAAGT GAGCATATGA AGTCCATTTT 720

ACATGGCTAG TACATATAAC TTTTAAAAAG TTGGACATAG TTATATTTTT CCATTTATTT 780

ATTTACTTTA TATCCTGATC ACAGACCCCC CCCTCCTCTG GATTAACTCT CTCCACTGCT 840

TCTTACCCCT CCCCATCTCT CCTTCACCTC TGAGAAGGGG GGATACCTCC TGTCTTATCT 900

GGTTTCAGTG GGAGAAGGAT GTATCCTAAC ACATATAATT TTTAATATCC TGAGTTTTTC 960

TTTCATACAC CTTACTTATT CTATTCATTT TTCAGGAAGG CATGTTTAAT GTTTTTTTTT 1020

TAATTTTATG TGTACGAGTG TTTTGCCTAC ACAGTCATAG TGCATCGCAT ACATTTTTGC 1080

TGCCCGTAGA GATCAGAAGG GAGCATTGGG TTCCCTAGGA CTGGAGGCAT GAACCACCTT 1140

GTGGGTGCAG AGAACTGAGC CTGGGTCATC TCAAAGCATC AGGTTCTTCT TGAGTCATCT 1200

CACTTGCCAC TTCTCCCATT TACTGATTTT ATCTGTGTGC AGACATTCAT GGCCCAGTCC 1260

ACAGGTGGAA GTCAGGGACA ACCTATAGGA GTCAGTCCTC TCCTTCTACC GTGTGAGTCC 1320

CTGGCCTCAA ACTCAGGTTG TCGGGCTTCA TAGCAAGAGC TTCTATTTGT TGAGCCATCT 1380

TGCTAGCCCC ACCCCATACT ATCTTTATAA TATCTGTTTA ATTAAGACAT TCATAATGAA 1440

TTTTATTAAC ATTCATCGTT ATCCCCTTTA CCAATTTTAC TATGTATTAA TTGCCACCCC 1500

TTTAAATTTA ATTACTTCCT TGGCTGGGTT TTACAGGAGA GTTCCAGGAA GCTAGATGGA 1560

GAGATGGCTC AACAGTTTAG AGCAACGGCT GTTCTTGCAG AGGACCTAGG TTCAAGTCCT 1620

GGCACTCAGA GGTGGCTCAC AATCATCTGT GACTTCAGTT CCAGGGGATC TGAAGAATTC 1680

TTCTGGGCTC CATGGGCATC AACTACACAC TTGGTTCATA GACATACATG CCAGCAAATG 1740

ATTGATCCAT ACATATGAAA TAAACCATAA ACAGAAAAAA AAAAGGAAGG TGAGGGAAGG 1800

AAAAAAAGTT TAAAAAAAGG AAAGGAAGGA AGGAAGGGAN NNNNNNNNNN NNNNNNNNNN 1860

NNNNNNNNNN NNNNNNNNNN NNNNTCTCTC CATACTGAAA GATGTCCACA ATGACTAAGG 1920

GAATTTTTTT TAAAAGACAA GCACAACGTT TTCTAGGGAT CAAACTCTAT TTGTGAGGAA 1980

GACTGGTGGT TTGAAGATTA CATAGCAGAG TTACATCTAA CATGAGCGTG TTTCCCCTGG 2040

ATGGAAGGAG TCTGATAACT TGTCTTTCTT TCTTAGTTAG CATCTCAGAG TCCCCCGCCT 2100
CCCTTAACAT CCTTTTTGCA CACCATCTTT TTAGGAAAAT GGATCATTTA TGGGGATGTA 2160
GTGATTTGTA CAAGAATGTC CCCTGTGGGC TCAGATATTT GAATACTTAG TTCCCAGTTG 2220
GGGGAGCTTT TGTAGGGAGG TTGGGAGGCA CAGCCTGGCA GGAGGAAGCA TGCTAGCAGC 2280
TTTGAGACTA TAAACCCTCA TCTACTACCT TGTTCTCTTT CTGCATTGTG CTGTGTCTGA 2340
CACTGTGAGA TTCCTGCTCC CGATGCCATG CCTGCCCGCC ATGATAGACT CCTAGCCCTC 2400
TGGAAAGGTA ACCTCAGTGA ACTCTCTTCT ATAAGTTTCT TTGCTCCTGG TGTTTTATCA 2460
CTGAAACGGA AAAGCTTGCA GGGAGGTAGG AGGCAGCCTG TGGCGTTGAT TCAATGCACC 2520
TGGCCTTATC CTCGGATGAG ATCGGTCACC AGTCAAAAAC TGTGAGCTTG AAGGTCTTGG 2580
GTGCTTAACA TCTATTTTTA CAAATCTTAT TTAGCAACTT AGAACTGTGA AATATTGGAA 2640
AGCTACTTAA ACCTTCTAAA CTCCCTCCTC CACACTATGA GAATGTTACA TTTTCTATTC 2700
AGTTATTTTT GAGCAGTAAA CAGATGAATC AAGGAATATG CCCATCACAT CAAGAGTGCT 2760
CCTAAATGGA CTTGCTTGTT ATTCATTTAC AGTGTGGCCC CTTGACTTTC ATCGGCACTC 2820
CTAGCAGAAA ACAAAATCCG CCAGATGGAG CTGGAGAGAT GGCTCAGCTG TTAAGAATAC 2880
TTATCCCTAC ACAGGCCCTG GAGCCAGTTC CCAGCACCCA CACGGTGGCT CACAACCATC 2940
TGTAACTCCA GTTCTAGGAG ACCCGACTCC CTCTTCTGTC TGAAAACACC AGGCACGCGT 3000
GCGGTCTACA TACAAACATG AAAGCAAAAT ACACACATTA CATAAATAAA TCTTAAAAAA 3060
TGATTCGGGG TGGGGGAAGG AAAAAAAAGG ATGTTAGAAA ATCGATGTAA CTGTTTTTTC 3120
CTTTTGCACA GATCTAAGTT AGGGAAGGAG AACATTCTCT TACCATCGAA AATAATTGTT 3180
TTCATTGCCC CCAAGTCTGC TAATAGAGCT TGCTACCTTC ATGGCTGTCG TAAGGATGAG 3240
GCAAAGATGG ACTTCAGCTT TCAGACTGTG TCTGCTCAAA TGTTGGCTAC TCCTGTTTTC 3300
TGACCCCCTT CTCTGGTGCA ATGTGGACTT TCAATTAATT TCCCTGCATC TTTTACATAT 3360
TTGATTTAAA AAATATTTTA TTTTATGTAA TTGTATGTAT ATGCATGTCA ATAAGCATAT 3420
GTGTGTGTGT TTCCATGGAA ACCAAGGCAA CAGATTTTCC AGAGCTGTAG AAATGGGCTG 3480
TGAGACGCCC ACTGTGGGTG TTCGGAACCA AACTCGGGTC CTGTGGAAAG ACAGCGAGCA 3540
CCCATAATGC AGAGGTATCT CTCAGATTTT ACTTTAAAAT TTCAATTTTC TTTTTTTTTT 3600
TTAAAGTTCC AAGTAACTAT AGGAAAGTAC ATGGGTATAT AGATCCCCAG TACCAAGATT 3660

CTTCCTTTGC AGGTAGCACA ACTTGGTTTG TTTCACATAA AGAATGGAAA GTCATTAAAA 3720

CACTCATCAC ACTGTAAAGT AGAATTGAAC TCTGACAGAA CAAGCGAAGT GAGTCTGACT 3780

TCCAGGTAAC TGAGCCTTCT TTTCCTCCTA AAGACACAAG CCATACACAG AGTAAAATAA 3840

ACTTGGGCAT GGTGAGAAGG AAACAACGCA GGAGGGCTAG CCAAGTCTGA GAGTCGTGAG 3900

TGTGCTCGGT TTATAAACGG AGCCCACCTT GCCAGCGAGG TAGTCACATG CTCTGCTAAA 3960

CAGAAACTTA AGAAAACACT TACACGAAGC AAACATGGGG AAGTGCCATG CAAGCATGTG 4020

ACTGACTGGT GGCAATGACC GAAACCACAG CAGCCACTAG AAAAGGAAGG GTAGTGCGCC 4080

ACACTGTAGT TGTGAAAATG AACTTATTCA TTTATTTTGA AAAACGTGTA AGAAGCAAAG 4140

ATGTCTTCTT TCCCACCTAC CTTTGCGGCA GGCGAGCACT TCCTGGAATT TATAAAGTGC 4200

GATCTTTCTG GGGACTTCTC ATAACATTTC CTACTGCTCA TCTATGTCTG TGTCAAATAG 4260

AGAATGCTCT TGAACAAGTG TGTGTGTGTG TGTGTGTGCG CGCGCACGCG CACTCACTCC 4320

TGCTCTGTTG AGGTCCAGTT TTGATGGTCC CGCCAGAGGT ATATTTGAGT ATCATTTCTC 4380

AAGAGCTTCA GCTGGGAGAC ACTGCCTCTT ACTGGCCTGA AGGTCACTAG CTGATTCATC 4440

TCCGTTTGGG CTGGCGCGCC TTGGGGATCC TCCTATCTCT CCTTCCCCAG TGCTGGGATA 4500

ACAAGGTTGG CACCACATGA GCCTTTTAAA ATGTGAGTTT GGAAGCTCAA ACGCAGGTTT 4560

TCATGCTTGC ACTGAAACTT CACAAGCTGA ACCGTCTCCC TCTCCTTCCC TCTCTTTTTT 4620

CCTTTTCTTC TTCCTTTTTA AAACACATCT TGTCTTTAAA AAAAAAAAAA GGCCCAAAAC 4680

AAGTGTAAAG TATTTCCCTA TGTGTGTGGA GGGAGGGAGT ATAGGAGGCT GATTTCACTG 4740

AGATCCTGTT AAATTTGGGT GCCATAGCCA ATCAAAGACG CATCGTTTCC TCTAAGAATT 4800

CTAAATGGGG CGATTACCAC GGGCCTGCAG GTTCTGGTTT GTATTAGAGG AGACACTGTC 4860

TTCTTAAGTA AAACATAGAA GGGGAAGTGT CCAGAATTGT AAATAAGGCT TCGAGAGAAG 4920

CCTTGTCTGG CCACCGGGAT GGAGAAGACC TACCTTCGCC TATCCAGGAT CCATCGTCCC 4980

TCCCTCTACC CAGATCTGAC AGCCCTCCTT GGCTCTTTTG CTGAGGTTTG TTTGAGTTTG 5040

TTTTACTCTC TGCAAGAGAA GTTTCCTTAA ACATTCTACC CTGTTCACAA GTAAATACAC 5100

CTCTTAGCTA AGAGGCCACA CACCCAGGGG GAACACCGAT AAAAAGAACA AGCCAGAACC 5160

TTCAGAACGC TGTCGATAGG TACACCAAGC AGCCTTCATA CGGAGTTTTC ATTCGTGAGG 5220

AGCTGAATAT ACAACAAAGC TAAATGTGAG CAGACCAGGC ATGCCTCTGC TAAATGAGGA 5280

TGCCCACACC AAACATGCCC AAGATCTTCA AGTATAATTT TATTATATAG ATTCGCTATG 5340

TGTTGACATG TTTTTATAGT GAACCTGGAT TTTACAAACC CTCCTGGTTT GCCACCTGCT 5400

TCTGGCACCA TACTTGAGGC TTAGGCACGT GATAAAGGAG CATGCCTGTT TCCCCCCTTA 5460

TTTTTTTTAA AGAAAAGCAC CATGTTACAT CATTAATCAT GCATATCAGT GTAGTTTAGA 5520

TCCGATGTAG AGACAATAAT CTTATCTCTT TGTCTGGCTG AAAGACTGTC CTTTAAACTA 5580

TCATTCTAAA TGCATTTGGT TTTTGCCAGG AGTAAAACAT GTCACAAGAT ATTTGTTGTC 5640

ATTTCCCAGG CGTGGAAGGA AAGGAATGGA AAGAAAACCA GGGGTGAAGG CTGCTGTTCC 5700

TCTCTAGTCG CTACTTGAAG TCTACATAGC TGGGGGGGGG GGGGGGACTG TTCACATGGG 5760

ACCGGTTTCC TCTTTGTTCC TACACTGGCG CCTCTGGCAA AAAACTCTCC CTTCTCTTCC 5820

CCCCAAGCAT ATCTTGGCTG AAAGGTCAGC TCTGAAAAGG GGCCTGGCCA AAGTTACTGT 5880

AGGGGACCGT GGTCATGGAA CTGGGTAAAC AAAAGCACTC TAGCAGCCAC TGGAAAAGGA 5940

CCGGGGGCTC TTCTCTGTGC ATTTGCCCTG GAACCCTGAC CACCGCCAGC TCCCTGCATC 6000

TCCTTGCTAT GGGTTTTCTG GACCGACCCA GCCAGGAAGT TCACAACCGA AATGTCTTCT 6060

AGGGCTAATC AGGTAACTTC GGACGATTTA AAGTTGCCAG ATGGACGAGA AAACAGTAGA 6120

GGCGTTGGCA ACCTGGATAA GCGCCTATCT TCTAATTAAA ACATTCAGAC GGGGCGGGGG 6180

ATGCGGTGGC CAAAGCACCA TAAAACAAAA CTTCCAAGTA CTGACCAACT CACTGCAAGT 6240

TTGTGCCCCG AGTACATCTA GGTTCAGGGG TTCTTGTCTT CATGCTCCCA ACTGCGGGCG 6300

GATTTTTGGT CCCTTGGGAC TTTCAGTGCA GCGGCGAAGA GAGTTCTGCA CTTGCAGGCT 6360

CCTAATGAGG GCGCAGTGGG CCTCGTGTTT CTGGTGATGC TTCCCAGGTT GCTGGGGGCA 6420

GCAAGTGTCT CAGAGCCCAT TACTGGCTAC ATTTTACTTC CACCAGAAAC CGAGCTGCGT 6480

CCAGATTTGC TCTCAGATGC GACTTGCCGC CCGGCACAGT TCCGGGGTAG TGGGGGAGTG 6540

GGCGTGGGAA ACCGGGAAAC CCAAACCTGG TATCCAGTGG GGGGCGTGGC CGGACGCAGG 6600

GAGTCCCCAC CCCTCCCGGT AATGACCCCG CCCCCATTCG CTAGTGTGTA GCCGGCGCTC 6660

TCTTTCTGCC CTGAGTCCTC AGGACCCCAA GAGAGTAAGC TGTGTTTCCT TAGATCGCGC 6720

GGACCGCTAC CCGGCAGGAC TGAAAGCCCA GACTGTGTCC CGCAGCCGGG ATAACCTGGC 6780

TGACCCGATT CCGCGGACAC CGCTGCAGCC GCGGCTGGAG CCAGGGCGCC GGTGCCCCGC 6840

GCTCTCCCCG GTCTTGCGCT GCGGGGGCGC ATACCGCCTC TGTGACTTCT TTGCGGGCCA 6900

GGGACGGAGA AGGAGTCTGT GCCTGAGAAC TGGGCTCTGT GCCCAGCGCG AGGTGCAGGA 6960

TGGAGAGCAA GGCGCTGCTA GCTGTCGCTC TGTGGTTCTG CGTGGAGACC CGAGCCGCCT 7020
CTGTGGGTAA GAAGCCCACT CTTTAGTAGT AAGGCGGAGA AGTAGGGTGC GGGCGGAGAG 7080
TGGGAATAGA AGAGGACCTA ACTCGTAGAG CTCTAGAGAC CCTCCTCCCT TGGGTGTTCT 7140
TTCACTTACC AATGGGGAAA CTGAGGTTCA AAGACTCTTC CGAAATGACT CAGCCAGGAT 7200
TCTACTCTCC CCCGGGCATC GGTTGGAGCG TGTCCTGCGG AGCCGTCACA GCCCCTGGCG 7260
CTAGGTAGGC AGGAGTGGAA AGGCGGCCTG AGCCGGGGCA GGAGATGCTC CCACTGGCAG 7320
GAACAGGCGG TCAAACGCTG GGAAGCCAGC TCAAGCCAAG CGGCCCGGCT GGCATCAATC 7380
ACTCGGTGCT GTTGCCCACC GCCCTAGTGG GGGGCAGGGA ATCCGCCTCT GGCTCCGCTC 7440
CCCTTTAGCT CCAGCGTGTA AGCGCACGGA CTATGTGAGG GTAGGTCTCT TCATAGAGCA 7500
ACACTTTCCT CCCTCAACTT TCTTTGATGC AGAATGCTAT TTTTGCTGGT AGGAGGAAGA 7560
CGCGGCTTTC TCTTCTGTGA CAGCTTCTCC AGGTGTATTA AACTAAATAA CTCTCCACTT 7620
ACCGACTCCA AAGCGCTGGT CCTGGGGTAA ACTCTGAAAG TCTCAGAAAC TCTTGAGCTT 7680
GGCACCTAGT TATAGGTCAC TTTTCTTGTT TTAAAATGCC CTCTGCTTCA AGGTTAGGCC 7740
CACACTCGCT CTTGGGCTTT TGTGCAATAA TTTCCCTTCC CTTCCCTTCC CTTCCCTTCC 7800
CTTCCCTTCC CTTCCCTTCC CTTCCCTTCC CTTCCCTTTC CCTCTTCCTT TTCCTCCTCC 7860
TCTTCCTCCT CTATTTCTCT GTCATTTCCT TTTTGAAGCC ACAGTTTGCA GATTTCCAAT 7920
CTCCACCCAT TGGAGAATGG AGAATCAGGA AAAAAGAAGT CAATTCTGCA GAAACATTCC 7980
TTGCGCCCTA AGAGAATCGC ATGGCTTAAA AGCATTGGCA CTGACATACG GCGCCAAGAT 8040
CGCCTGTCTA GAGCTATTGA GTTTTCCTCA TAATGACTTG GTTCATCAGG CTAGCTCCAC 8100
CACGAGTGCC CTCTTGTTCC TGAGAAGGCC GCACTCTCCC CCTTTCTGGG AAGAGAAAGA 8160
CAGCCTGGAA CATGTGCTTG CCCTGGGTTC CATAGAGAAG CAAGTTGCTT TAAAGCCCAG 8220
AGAATTCCTA GTGTAGCAGC TTAACAGCGT CCCGTTCTCT GAATAAGATG GAGGTTGCCC 8280
TTTTGGAGTG TGTGACTTGC TTAATTGGAT TGGGCTATAA TTGGTGCCAT CCAAGTCTCG 8340
AGACAGAGCC GCTGTTGTTT TTCCTTCTGG TCTTTGAGCG GGAAGGATAA CAGTGCACAA 8400
ATTAATTAAT GTTGGTTATC GGATTTGAAC ATAAAAGGGC TTTTATTGTA TAGTAGCATA 8460
TGTACCTCTT GCAGTCAGAA TGAGCTGTCT AAAGAACAGA ACCCAAACTT GCCGATGAAA 8520
ATGAATGAGG TTTAATAAAG GCGATGGATG AGCATTAGTC ACTGATGTAA ATCTCCAGTT 8580
ATTGATAACC TCATTGACTG GATTTGATTG CAGACATGTA TTGGTATGGG GCATCCTTTA 8640

AAGATGAGCA TAGCCAACGT GCCTGCACTC TAAGAGAATC TATGGCTGTA TGTTATTACA 8700

GAGACAGTTG AGAAGCTCTT AGTGGCTCTG GCGTGTAGAT CAGCGGTAGA GCGCTGAGGC 8760

TCTGCGCTCG CTTCCTGGCA CTGAAGAATA AAGGCCATTT ACTGTGGTGG TGCAGTGGGC 8820

GCAGTTTGTG ACGAGTTACT ACTACATTTT CCTCACACAT CTGCCTGACT AATGAGTTCA 8880

TCAGATGAGC GTATCCAGTG ATTGTTTGCA GGTTAATGGT TCTCAGTCAT GTTTAGAATC 8940

TACTTATCAA ACAAATTGTT TTCTCATTTC CTGCTTCTTC TCAAACAAAG TAAGATTCCA 9000

TTATTGAAAG GCTTGTTTAA GAGCATTTTA ACTGCTTGCC TATGTTAGGG ACAGTGACTT 9060

ATTTCATATT GACAAATATT ATGCCGATTA ATTGAATATG ACTACCCAGT TCTATAGCTG 9120

TCTCAGGGCA GACCAAGAGC ATCTGTGATC CAGTCACTTT AAATGCCATT TAAAATGCAT 9180

AATTTGTTGG TCTAGGAATA AACACACTGT AAAGTTTAGA ATCACGGCCC AAACACAAGT 9240

CTTTAACAAT GCCAACTAGC TTCTGAGATT CATTAATGTC ATTTAATTAC CAATGTTTTA 9300

AAAATATGTC ATTAATTACT AAATCTATAG TTGTAACAGC AACACATGTA CATCTTATTA 9360

AGTTGGGTAT ATTCAGGGTG GCATAGCTGT AGACTATTGC ACATCTGTGT TGGTGAGCCA 9420

GTGGAGAACT GCCTCCTGGC TGTTCTCAGA AGGCCACAGT GTCACGGCAT TGGCTATTTG 9480

CCTTGGCTCT TTGCTAATAC TTTATTGACA TGGCCTCATC TTCGTTCACG TTCACTTATT 9540

TGCCCAACAA CGTCAATGCC AGCTGAGGCC TTAGGAGTCA TCTGTTCTTA GTCAGTGCGA 9600

ATTAGAAAGC CTGGATGCCT GCCTGCTATT AATTAGTTAT TCTTCTCTTC TGAGACAGAG 9660

TCTCACTGTG TGGCCCAGGC TAGTCTCAAA CTTGCGGTCC ATTTGTCTCA CTCATCAGAA 9720

TGCTGGGCTT CCAGGTGTGT GCACCACACT AGGTAGCTCG CGTTTTAAGC TAAGAGCTGG 9780

AAGATCCTGA TGTCCTTTAC CATGGTGGGC ATGTTACAGG TTAGTTGACT GAAAACTAGT 9840

TATCTCGCTG TGTAATGACC TGCAGTGGTA TGTATCTCTC AAGATGCTTT TTTGCATTTC 9900

AATCAGTTAG GTAACAAGTT CTTAAGTCTC CAGCTTGGTA TTGGCATGAG CTCAGAGCTT 9960

TGATTAATGA GTTGGGACCC CCTAGCTATT GCTCATTAGA CTTACACTAT TTTTAGTTTT 10020

GCTCTGAGTT TATGAATATG CATGTATGCA TGAACTTGGG AGATATTTTT CTTCCCCAAT 10080

TCCTTTTCCT CCATTTAAAT GTGCTGTCTT TAGAAGCCAC TGCCTCAGCT TCTGCAGCTC 10140

AGATACCAAA GGAAGTCTGG TACACAGCAT GATAAAAGAC AATGGGACGG GGTCACAGTG 10200

GCTCCCGTCC CTTTCAGGGG TATGGAGACG AGCTGTAGAG AGATGTCTCC AGGGAGTTTT 10260

CATTAATCAG CAATTTAGTC AGATCTGTGC ATCCTATGCT TTACAAGAAA TGTCAGTGGG 10320

CCTGAGATCA TCAGATGGAG GTTCATCGGG TTTCAATGTC CCGTATCCTT TTGTAAGACC 10380

TTGAAGTTGG CAACGCAGGA AAACAGGAAC TCCACCCTGG TGCCGTGAAT TGCAGAGCTG 10440

TTGTGTTGGT TTGTGACCAT CTGCCCATTC TTCCTGTTAT GACAGAGCTT GTGAACTTTA 10500

ACTGGGACTG GGGCAAAGTC AATCCCACCT TTATACAATG AATTGCTGAA GAGGCCTTTT 10560

AAAACTTGGA GTGTGCATTG TTTATGGAAG GGCTTTCCTA TTGGATCCAA CTCTTTTCTA 10620

ATTTGTTTCT AGGTTTGCCT GGCGATTTTC TCCATCCCCC CAAGCTCAGC ACACAGAAAG 10680

ACATACTGAC AATTTTGGCA AATACAACCC TTCAGATTAC TTGCAGGTAA GGATTCCTTT 10740

TTGAGCCAGC TTTCCTATGT GAAAGGACTC ATTGTTTACT GAGGTCACAA CAATTTCCAC 10800

TATTGCAGAA GTATAATAGT ATTGTTACAA TTGTTTATAA ATCATGAGAC TTCTAAGAAC 10860

CTATTTAATA ATGAAACAAT GGAAAAAGTC TTTTCAAACC TTTGTACTCT TTTGCTGAGC 10920

CGTTTTCAAC ATGCACAAAC ATATTACACA AATATAACAT ACACAGGAAC ACACATGAAT 10980

GCATGGGATG ATGTGCCTAA AACTAGCATG TAATTGATAT TCACAATTAT TGATAAATTA 11040

GTAAAGCAAA GGAATTCCTT ATGAATAGAG CTAAAATTCT ATCCATGTTC AAGTCACCCA 11100

GAATGGCTTC TGGACATTTT TTTTTTTAGC TGTTTTCTAC AAGTGAAATT CTGCCTGTAT 11160

TAGCAATTTA ATATCTAGCC AATAATATTC CTGACCATAT GTCCTGTTCA GACCATGACC 11220

TTCATAATCT GGCTTGATGT TCTGGGCTTC TTTCCCTCTT GCCAGCAAGA TGTCACGGTG 11280

TTGATGCTGG ATAAACTGAG AAACAGAAGT TTTTCGCAAG AAGAGGACCT TGAATTTTGC 11340

TTTTCCCCTG AGAGACAAGA AAGGAAACTT AGAGGAGGTG TAGCTGGGAG TGTGGTCATT 11400

CATGAAAGAC CTGTTTGCAG GGCAGTGTGT TTTGCTGGGG ACAGTAATGA GCCTAGATCG 11460

TAGTGCCATC CCAAGAGAGT GCTTGGTGGC AAAAAGAGCC CTAGCAGCTT GTGGCAGTTG 11520

CCTCATATTT GAAGAATACT AAGAGGTCCC CCGAATAACT CAGGGCTAGT GTTGATCATT 11580

GCATGTGGAG AGAATCCAAG CCTCCTATCT AGGGTCTACA AAAGTAACCA ATGCCCAGTC 11640

TTTGGGGGAA AGCAAAACCA GAAAGCGATG ATAGCAGGAC CTGTTTATTT TCATTAAGTC 11700

ATGGCATTTC CAGAGACTTT GCTCCCCCTA TTCTCAGACA CAAAGCCCAC TTAAGATCTC 11760

CCTCTGGAGA CTGCTGGGAA CATTTCTTAA GTTCTGAAAA AACCCTGGAG TGATTGGGCA 11820

CAGACGATCC TGTCACTTCA TGTGAGTGCT AAGCTCTTTG GGTGATGACT CAGTGGGTCA 11880

CATTGTTTTA TTCATATTGA CTACCTTCCG TTTGCTTTGC GGAGAATGGA AGCTATAGAA 11940

GTCTGTTTGG TGTGGCCCTC ACAAGGCACT GTGAGCTTCT TCTCTCTGTG TGCTAACTTC 12000

TTACTCTCCC TTGCTTATAC CCACATAGGG ACTCTGGCTT TGTTGCTGTT CTTCAATGCT 12060

TCAGATGTGC CCTGGGTCCT GTCTGTCCTT CACACTTACT GATGCTGCCT GGAATGCTAT 12120

TCCTCCCAAT GTGCATAGGG CCAGCTCGGT CCAAATCCTC TCTTTTCTTT GCCTCTTTTA 12180

TATTTTCCTT CACAGTATCA AATCACCACA GTTTATGCAA CAAACTGAAA CTTTAAAATT 12240

GTCTGTCTCC TTATATTAGT GATAGGTTCC AGAAAGGCAC TGATTTTTTT TCTTCCCTGG 12300

TGTACACTGG GCAACTACTC TACCACTGAG CGTGATATCC TTGGTCCCTT AAAAGTTATC 12360

CTCTGTCCTT AATAATGCTT AGCAATCATA TTTGCTTAAA ATATTTATTG AATGACTGCA 12420

GGAATGAATG AATGAATGAG CTAACAGAAA ACTCATGACC ATGTGGGTGA TTTCCGAAAC 12480

AGAGTGTGAG ATCTTTGGTG GCATGTCCTT GTAGACTGTC TGCCACCAGT ATCTATCATC 12540

TTGAAGGTGA CTATTGAGTA GTTTATATGC ATGTGAAAAA CCAAACCTTC TATTCTCTTA 12600

CTCATAGCCT CTCTTAATCA TAGCCCTGTG GCATGGAGTG TACCATTGAT ATCTTCCTGG 12660

AATACTTTTT CAGGGGACAG CGGGACCTGG ACTGGCTTTG GCCCAATGCT CAGCGTGATT 12720

CTGAGGAAAG GGTATTGGTG ACTGAATGCG GCGGTGGTGA CAGTATCTTC TGCAAAACAC 12780

TCACCATTCC CAGGGTGGTT GGAAATGATA CTGGAGCCTA CAAGTGCTCG TACCGGGACG 12840

TCGAC 12845

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GGGGTACCGA ATTCTAAATG GGCGATTAC C 31

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GTGGTACCCA AACACTCAAC ACCACTG 27

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

TCGGTACCGA CCCAGCCAGG AAGTTC 26

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

TTGCTAAGCT TCCTGCACCT CGCGCTGGG 29

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

AGGGATCCAC TCTTTAGTAG TAAGGCG 27

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

ACCTCGAGAC TTGGATGGCA C 21

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GGGCTATAAT TGGTGCCATC C 21

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GGATGGAGAA AATCGCCAGG C 21

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

GTGTGCATTG TTTATGGAAG GG 22

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CATAGACATA AACAGTGGAG GC 22

4

B 1730 EP

Claims

1. A recombinant DNA molecule comprising:
   (a) at least one first regulatory sequence of an intron of the Vascular Endothelial Growth factor (VEGF) receptor-2 (Flk-1) gene or of a gene homologous to the Flk-1 gene being capable of conferring expression in endothelial cells in vivo; and
   (b) operatively linked thereto a heterologous DNA sequence.

2. The recombinant DNA molecule of claim 1, wherein said first regulatory sequence comprises a GATA-binding site, an AP-1 binding site, an SP1 binding site, a PEA3 consensus sequence or any combination(s) thereof.

3. The recombinant DNA molecule of claim 1 or 2, wherein said first regulatory sequence is selected from the group consisting of
   (a) DNA sequences comprising a nucleotide sequence as given in SEQ ID NO: 1, preferably from nucleotide 8260 to nucleotide 10560;
   (b) DNA sequences comprising the nucleotide sequence of the human Flk-1-intron;
   (c) DNA sequences comprising a nucleotide sequence which hybridizes with a nucleotide sequence of (a) or (b) under stringent conditions;
   (d) DNA sequences comprising a nucleotide sequence which is conserved in the nucleotide sequences of (a) and (b); and
   (e) DNA sequences comprising a fragment, analogue or derivative of a nucleotide sequence of any one of (a) to (d) capable of conferring expression in endothelial cells.

4. The recombinant DNA molecule of any one of claims 1 to 3, wherein said heterologous DNA sequence is operatively linked to further regulatory sequences.

5. The recombinant DNA molecule of claim 4, wherein said further regulatory sequence is a promoter.

6. The recombinant DNA molecule of claim 4 or 5, wherein said further regulatory sequence is a 3'-untranslated region.

7. The recombinant DNA molecule of claim 5 or 6, wherein said promoter is a promoter of hypoxia inducible genes, genes encoding growth factors or its receptors or glycolytic enzymes.

8. The recombinant DNA molecule of claim 7, wherein said growth factor is VEGF, PDGF or Fibroblast growth factor.

9. The recombinant DNA molecule of any one of claims 5 to 8, wherein said promoter comprises a DNA sequence selected from the group consisting of
   (a) DNA sequences comprising the nucleotide sequence as given in SEQ ID NO:1 from nucleotide 6036 to nucleotide 6959;
   (b) DNA sequences comprising the nucleotide sequence of the human Flk-1/KDR promoter;
   (c) DNA sequences comprising a nucleotide sequence which hybridizes with a nucleotide sequence of (a) or (b) under stringent conditions;
   (d) DNA sequences comprising a nucleotide sequence which is conserved in the nucleotide sequences of (a) and (b); and
   (e) DNA sequences comprising a fragment, analogue or derivative of a nucleotide sequence of any one of (a) to (d).

10. The recombinant DNA molecule of any one of claims 1 to 9, wherein at least one of said DNA sequences is of human or murine origin.

11. The recombinant DNA molecule of any one of claims 1 to 10, wherein said heterologous DNA sequence being operatively linked to said regulatory sequences is located 5' to said first regulatory sequence.

12. The recombinant DNA molecule of any one of claims 1 to 11, wherein said heterologous DNA sequence encodes a peptide, protein, antisense RNA, sense RNA and/or ribozyme.

13. The recombinant DNA molecule of claim 12, wherein said protein is selected from the group consisting of Vascular Endothelial Growth Factor (VEGF), Hypoxia Inducible Factors (HIF), HIF-Related Factor (HRF), tissue plasminogen activator, p21 cell cycle inhibitor, nitric oxide synthase, interferon-γ, atrial natriuretic polypeptide and monocyte chemotactic proteins.

14. The recombinant DNA molecule of claim 12, wherein said protein is a scorable marker, preferably luciferase, green fluorescent protein or lacZ.

15. The recombinant DNA molecule of claim 12, wherein said antisense RNA or said ribozyme are directed against a gene involved in vasculogenesis and/or angiogenesis and/or tumors of endothelial origin.

16. A nucleic acid molecule of at least 15 nucleotides in length hybridizing specifically with a recombinant DNA molecule of any one of claims 1 to 15.

17. A vector comprising a recombinant DNA molecule of any one of claims 1 to 15.

18. The vector of claim 17, which is an expression vector and/or a targeting vector.

19. A cell transformed with a DNA molecule of any one of claims 1 to 15 or the vector of claim 17 or 18.

20. The cell of claim 19, which is a prokaryotic or eukaryotic cell, preferably an endothelial cell.

21. A pharmaceutical composition comprising a recombinant DNA molecule of any one of claims 1 to 15, the vector of claim 17 or 18 and/or the nucleic acid molecule of claim 16 and optionally a pharmaceutically acceptable carrier.

22. A diagnostic composition comprising a recombinant DNA molecule of any one of claims 1 to 15, the vector of claim 17 or 18 and/or the nucleic acid molecule of claim 16, and optionally suitable means for detection.

23. A method for the production of a transgenic non-human animal, preferably mouse, comprising introduction of a recombinant DNA molecule of any one of claims 1 to 15 or a vector of claim 17 or 18 into a germ cell, an embryonic cell or an egg cell or a cell derived therefrom.

24. A transgenic non-human animal, preferably mouse, comprising stably integrated into its genome a recombinant DNA molecule of any one of claims 1 to 15 and/or the vector of claim 17 or 18 or obtained according to the method of claim 23.

25. A method for the identification of a chemical and/or biological substance capable of suppressing the transcription of a gene in endothelial cells comprising:
    (a) contacting a cell of claim 19 or 20 or the transgenic non-human animal of claim 24 either of which is capable of expressing the heterologous DNA sequence with a plurality of compounds; and

(b) determining those compounds which suppress the expression of said heterologous DNA sequence.

26. A method for the identification of a chemical and/or biological substance capable of activating and/or enhancing the transcription of a gene in endothelial cells comprising:
    (a) contacting a cell of any one of claim 19 or 20 or the transgenic non-human animal of claim 24 either of which is capable of expressing the heterologous DNA sequence with a plurality of compounds; and
    (b) determining those compounds which are capable of activating and/or enhancing the expression of said heterologous DNA sequence.

27. Use of a recombinant DNA molecule of any one of claims 1 to 15, the vector of claim 17 or 18, the cell of claims 19 or 20, the pharmaceutical composition of claim 22, the diagnostic composition of claim 23 and/or the transgenic non-human animal of claim 24 for the identification of a chemical and/or biological substance capable of suppressing or activating and/or enhancing the transcription, expression and/or activity of genes and/or its expression products in endothelial cells.

28. The method of claim 25 or 26 or the use of claim 27, wherein the chemical and/or biological substance is selected from the group consisting of peptides, proteins, nucleic acids, antibodies, small organic compounds, hormones, neurotransmitters, peptidomimics and PNAs.

29. A method of inhibiting a vascular disease in a subject, comprising contacting an artery of said mammal with the vector of claim 17 or 18, wherein said protein reduces or prevents the development of the vascular disease.

30. The method of claim 29, wherein said protein reduces proliferation of smooth muscle cells.

31. Use of a recombinant DNA molecule of any one of claims 1 to 15, the vector of claim 17 or 18, the nucleic acid molecule of claim 16 and/or a substance identified by the method of claims 25, 26 or 28 for the preparation of a composition for directing or preventing expression of genes specifically in endothelial cells.

32. Use of a recombinant DNA molecule of any one of claims 1 to 15, the vector of claim 17 or 18, the nucleic acid molecule of claim 16 and/or a substance identified by the method of claims 25, 26 or 28 for the preparation of a pharmaceutical composition for treating, preventing and/or delaying a vascular disease and/or a tumorous disease in a subject.

33. Use of a recombinant DNA molecule of any one of claims 1 to 15, the vector of claim 17 or 18 and/or the nucleic acid molecule of claim 16 for the preparation of a pharmaceutical composition for inducing a vascular disease in a non-human animal or in the transgenic non-human animal of claim 24.

34. The method of claim 29 or 30 or the use of any one of claims 31 to 33, wherein the vascular disease is atherosclerosis and/or a neuronal disorder.

35. Use of a regulatory sequence as defined in any one of claims 1 to 3 for enhancing and/or directing gene expression in endothelial cells.

5

Abstract

Described are recombinant DNA molecules comprising the regulatory sequence(s) of an intron of the Endothelial Growth Factor (VEGF) receptor-2 gene (Flk-1) or of a gene homologous to the Flk-1 gene, being capable of conferring expression of a heterologous DNA sequence in endothelial cells, preferably in vivo. Vectors comprising said DNA molecules are provided. Also provided are pharmaceutical and diagnostic compositions comprising such recombinant DNA molecules and vectors. Furthermore, cells and transgenic non-human animals, comprising the aforementioned recombinant DNA molecules or vectors stably integrated into their genome and their use for the identification of substances capable of suppressing or activating transcription of a gene in endothelial cells are described. Described is further the present invention also relates to the use of the before described recombinant DNA molecules and vectors for the preparation of pharmaceutical compositions for treating, preventing, and/or delaying a vascular or tumorous disease in a subject. Furthermore, the uses of the recombinant DNA molecules and vectors of the invention for the preparation of pharmaceutical compositions for inducing a vascular or tumorous disease in a non-human animal are provided.

Figure 1

```
-6660 TCTAGAATAT AGAAGATAAG TTTGCGTACA ATTCAGTCCT TTGAAGACCT
      GATAAGCTTT AAGAAGGAAG ATGGGTTACA CATTGGGAAA TGGTTGCAAT
      CTGCACATGG CAGAGGCAAG AGATGCAAAT CACATTTCTT ACATACTCCA
-6510 TACAAATCTT ACAAGACTGT TTTTCTTTCT CATTTAAAAT AAGAAGACCT
      GCCAGTCTTC CCCTTATTAC TAATTACAGT CACTCTGTAT CTTTGTTGAC
      ATTGGATAGT TTTACATACT TCAACAGGCT GGTGTCATTA AAGTTGTGGT
      GGGTGGGCAC CAGAGACACG TGATTCAGAG TGGGAGGAGA TGCAGGAGAA
      ACGAGGCACA GCAGAAGCAG AAGCGAGGAA AAACACTCTC AACGTTACTA
      ACACATCGAG AGGTTCCGCA CACTAGCAAT ACGGGCTGAA TCTGACCTAA
      TCTCTGCTGT TGAAAATTTT GCCTAGCCGC ACACTAGCAA TACGGGCTGA
      ATCTGACCTA ATCTCTGCTG TTGAAAATTT TGCCTAGCCT GTCACACAAG
      TGCTGAGCAT ACAGAAAAAG GAGAGTAATT CTCTGGTTCT TTGACTAACC
      AAATAGTCTA TATCAAATTG CCTAAGATAA TGTATACATT TAGTACATGA
-6010 CTGGTTATAC CTATTCTATA TGACTATTAT TTAAATGTGA ATTTACAAGT
      GAGCATATGA AGTCCATTTT ACATGGCTAG TACATATAAC TTTTAAAAAG
      TTGGACATAG TTATATTTTT CCATTTATTT ATTTACTTTA TATCCTGATC
      ACAGACCCCC CCCTCCTCTG GATTAACTCT CTCCACTGCT TCTTACCCCT
      CCCCATCTCT CCTTCACCTC TGAGAAGGGG GGATACCTCC TGTCTTATCT
      GGTTTCAGTG GGAGAAGGAT GTATCCTAAC ACATATAATT TTTAATATCC
      TGAGTTTTTC TTTCATACAC CTTACTTATT CTATTCATTT TTCAGGAAGG
      CATGTTTAAT GTTTTTTTTT TAATTTTATG TGTACGAGTG TTTTGCCTAC
      ACAGTCATAG TGCATCGCAT ACATTTTTGC TGCCCGTAGA GATCAGAAGG
      GAGCATTGGG TTCCCTAGGA CTGGAGGCAT GAACCACCTT GTGGGTGCAG
      AGAACTGAGC CTGGGTCATC TCAAAGCATC AGGTTCTTCT TGAGTCATCT
      CACTTGCCAC TTCTCCCATT TACTGATTTT ATCTGTGTGC AGACATTCAT
      GGCCCAGTCC ACAGGTGGAA GTCAGGGACA ACCTATAGGA GTCAGTCCTC
      TCCTTCTACC GTGTGAGTCC CTGGCCTCAA ACTCAGGTTG TCGGGCTTCA
```

TAGCAAGAGC TTCTATTTGT TGAGCCATCT TGCTAGCCCC ACCCCATACT

ATCTTTATAA TATCTGTTTA ATTAAGACAT TCATAATGAA TTTTATTAAC

ATTCATCGTT ATCCCCTTTA CCAATTTTAC TATGTATTAA TTGCCACCCC

TTTAAATTTA ATTACTTCCT TGGCTGGGTT TTACAGGAGA GTTCCAGGAA

GCTAGATGGA GAGATGGCTC AACAGTTTAG AGCAACGGCT GTTCTTGCAG

AGGACCTAGG TTCAAGTCCT GGCACTCAGA GGTGGCTCAC AATCATCTGT

-5010 GACTTCAGTT CCAGGGGATC TGAAGAATTC TTCTGGGCTC CATGGGCATC

AACTACACAC TTGGTTCATA GACATACATG CCAGCAAATG ATTGATCCAT

ACATATGAAA TAAACCATAA ACAGAAAAAA AAAAGGAAGG TGAGGGAAGG

AAAAAAAGTT TAAAAAAAGG AAAGGAAGGA AGGAAGGGAN NNNNNNNNNN

NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNTCTCTC CATACTGAAA

GATGTCCACA ATGACTAAGG GAATTTTTTT TAAAAGACAA GCACAACGTT

TTCTAGGGAT CAAACTCTAT TTGTGAGGAA GACTGGTGGT TTGAAGATTA

CATAGCAGAG TTACATCTAA CATGAGCGTG TTTCCCCTGG ATGGAAGGAG

TCTGATAACT TGTCTTTCTT TCTTAGTTAG CATCTCAGAG TCCCCCGCCT

CCCTTAACAT CCTTTTTGCA CACCATCTTT TTAGGAAAAT GGATCATTTA

TGGGGATGTA GTGATTTGTA CAAGAATGTC CCCTGTGGGC TCAGATATTT

GAATACTTAG TTCCCAGTTG GGGGAGCTTT TGTAGGGAGG TTGGGAGGCA

CAGCCTGGCA GGAGGAAGCA TGCTAGCAGC TTTGAGACTA TAAACCCTCA

TCTACTACCT TGTTCTCTTT CTGCATTGTG CTGTGTCTGA CACTGTGAGA

TTCCTGCTCC CGATGCCATG CCTGCCCGCC ATGATAGACT CCTAGCCCTC

TGGAAAGGTA ACCTCAGTGA ACTCTCTTCT ATAAGTTTCT TTGCTCCTGG

HindIII (-4200)

TGTTTTATCA CTGAAACGGA AAAGCTTGCA GGGAGGTAGG AGGCAGCCTG

Figure 1 continued

BstEII (-4100)

TGGCGTTGAT TCAATGCACC TGGCCTTATC CTCGGATGAG ATCGGTCACC

AGTCAAAAAC TGTGAGCTTG AAGGTCTTGG GTGCTTAACA TCTATTTTTA

CAAATCTTAT TTAGCAACTT AGAACTGTGA AATATTGGAA AGCTACTTAA

-4010 ACCTTCTAAA CTCCCTCCTC CACACTATGA GAATGTTACA TTTTCTATTC

AGTTATTTTT GAGCAGTAAA CAGATGAATC AAGGAATATG CCCATCACAT

CAAGAGTGCT CCTAAATGGA CTTGCTTGTT ATTCATTTAC AGTGTGGCCC

CTTGACTTTC ATCGGCACTC CTAGCAGAAA ACAAAATCCG CCAGATGGAG

CTGGAGAGAT GGCTCAGCTG TTAAGAATAC TTATCCCTAC ACAGGCCCTG

GAGCCAGTTC CCAGCACCCA CACGGTGGCT CACAACCATC TGTAACTCCA

GTTCTAGGAG ACCCGACTCC CTCTTCTGTC TGAAAACACC AGGCACGCGT

GCGGTCTACA TACAAACATG AAAGCAAAAT ACACACATTA CATAAATAAA

TCTTAAAAAA TGATTCGGGG TGGGGGAAGG AAAAAAAAGG ATGTTAGAAA

ATCGATGTAA CTGTTTTTTC CTTTTGCACA GATCTAAGTT AGGGAAGGAG

AACATTCTCT TACCATCGAA AATAATTGTT TTCATTGCCC CCAAGTCTGC

TAATAGAGCT TGCTACCTTC ATGGCTGTCG TAAGGATGAG GCAAAGATGG

ACTTCAGCTT TCAGACTGTG TCTGCTCAAA TGTTGGCTAC TCCTGTTTTC

TGACCCCCTT CTCTGGTGCA ATGTGGACTT TCAATTAATT TCCCTGCATC

TTTTACATAT TTGATTTAAA AAATATTTTA TTTTATGTAA TTGTATGTAT

ATGCATGTCA ATAAGCATAT GTGTGTGTGT TTCCATGGAA ACCAAGGCAA

CAGATTTTCC AGAGCTGTAG AAATGGGCTG TGAGACGCCC ACTGTGGGTG

TTCGGAACCA AACTCGGGTC CTGTGGAAAG ACAGCGAGCA CCCATAATGC

AGAGGTATCT CTCAGATTTT ACTTTAAAAT TTCAATTTTC TTTTTTTTTT

TTAAAGTTCC AAGTAACTAT AGGAAAGTAC ATGGGTATAT AGATCCCCAG

-3010 TACCAAGATT CTTCCTTTGC AGGTAGCACA ACTTGGTTTG TTTCACATAA

AGAATGGAAA GTCATTAAAA CACTCATCAC ACTGTAAAGT AGAATTGAAC

TCTGACAGAA CAAGCGAAGT GAGTCTGACT TCCAGGTAAC TGAGCCTTCT

Figure 1 continued

TTTCCTCCTA AAGACACAAG CCATACACAG AGTAAAATAA ACTTGGGCAT

GGTGAGAAGG AAACAACGCA GGAGGGCTAG CCAAGTCTGA GAGTCGTGAG

TGTGCTCGGT TTATAAACGG AGCCCACCTT GCCAGCGAGG TAGTCACATG

CTCTGCTAAA CAGAAACTTA AGAAAACACT TACACGAAGC AAACATGGGG

AAGTGCCATG CAAGCATGTG ACTGACTGGT GGCAATGACC GAAACCACAG

CAGCCACTAG AAAAGGAAGG GTAGTGCGCC ACACTGTAGT TGTGAAAATG

AACTTATTCA TTTATTTTGA AAAACGTGTA AGAAGCAAAG ATGTCTTCTT

TCCCACCTAC CTTTGCGGCA GGCGAGCACT TCCTGGAATT TATAAAGTGC

GATCTTTCTG GGGACTTCTC ATAACATTTC CTACTGCTCA TCTATGTCTG

TGTCAAATAG AGAATGCTCT TGAACAAGTG TGTGTGTGTG TGTGTGTGCG

CGCGCACGCG CACTCACTCC TGCTCTGTTG AGGTCCAGTT TTGATGGTCC

CGCCAGAGGT ATATTTGAGT ATCATTTCTC AAGAGCTTCA GCTGGGAGAC

ACTGCCTCTT ACTGGCCTGA AGGTCACTAG CTGATTCATC TCCGTTTGGG

CTGGCGCGCC TTGGGGATCC TCCTATCTCT CCTTCCCCAG TGCTGGGATA

ACAAGGTTGG CACCACATGA GCCTTTTAAA ATGTGAGTTT GGAAGCTCAA

ACGCAGGTTT TCATGCTTGC ACTGAAACTT CACAAGCTGA ACCGTCTCCC

TCTCCTTCCC TCTCTTTTTT CCTTTTCTTC TTCCTTTTTA AAACACATCT

-2010 TGTCTTTAAA AAAAAAAAAA GGCCCAAAAC AAGTGTAAAG TATTTCCCTA

TGTGTGTGGA GGGAGGGAGT ATAGGAGGCT GATTTCACTG AGATCCTGTT

AAATTTGGGT GCCATAGCCA ATCAAAGACG CATCGTTTCC TCTAAGAATT

CTAAATGGGG CGATTACCAC GGGCCTGCAG GTTCTGGTTT GTATTAGAGG

AGACACTGTC TTCTTAAGTA AAACATAGAA GGGGAAGTGT CCAGAATTGT

AAATAAGGCT TCGAGAGAAG CCTTGTCTGG CCACCGGGAT GGAGAAGACC

TACCTTCGCC TATCCAGGAT CCATCGTCCC TCCCTCTACC CAGATCTGAC

AGCCCTCCTT GGCTCTTTTG CTGAGGTTTG TTTGAGTTTG TTTTACTCTC

TGCAAGAGAA GTTTCCTTAA ACATTCTACC CTGTTCACAA GTAAATACAC

CTCTTAGCTA AGAGGCCACA CACCCAGGGG GAACACCGAT AAAAAGAACA

Figure 1 continued

AGCCAGAACC TTCAGAACGC TGTCGATAGG TACACCAAGC AGCCTTCATA

CGGAGTTTTC ATTCGTGAGG AGCTGAATAT ACAACAAAGC TAAATGTGAG

CAGACCAGGC ATGCCTCTGC TAAATGAGGA TGCCCACACC AAACATGCCC

AAGATCTTCA AGTATAATTT TATTATATAG ATTCGCTATG TGTTGACATG

TTTTTATAGT GAACCTGGAT TTTACAAACC CTCCTGGTTT GCCACCTGCT

TCTGGCACCA TACTTGAGGC TTAGGCACGT GATAAAGGAG CATGCCTGTT

TCCCCCCTTA TTTTTTTTAA AGAAAAGCAC CATGTTACAT CATTAATCAT

GCATATCAGT GTAGTTTAGA TCCGATGTAG AGACAATAAT CTTATCTCTT

TGTCTGGCTG AAAGACTGTC CTTTAAACTA TCATTCTAAA TGCATTTGGT

TTTTGCCAGG AGTAAAACAT GTCACAAGAT ATTTGTTGTC ATTTCCCAGG

-1010 CGTGGAAGGA AAGGAATGGA AAGAAAACCA GGGGTGAAGG CTGCTGTTCC

TCTCTAGTCG CTACTTGAAG TCTACATAGC TGGGGGGGGG GGGGGGACTG

TTCACATGGG ACCGGTTTCC TCTTTGTTCC TACACTGGCG CCTCTGGCAA

AAAACTCTCC CTTCTCTTCC CCCCAAGCAT ATCTTGGCTG AAAGGTCAGC

TCTGAAAAGG GGCCTGGCCA AAGTTACTGT AGGGGACCGT GGTCATGGAA

CTGGGTAAAC AAAAGCACTC TAGCAGCCAC TGGAAAAGGA CCGGGGGCTC

TTCTCTGTGC ATTTGCCCTG GAACCCTGAC CACCGCCAGC TCCCTGCATC

TCCTTGCTAT GGGTTTTCTG GACCGACCCA GCCAGGAAGT TCACAACCGA

AATGTCTTCT AGGGCTAATC AGGTAACTTC GGACGATTTA AAGTTGCCAG

ATGGACGAGA AAACAGTAGA GGCGTTGGCA ACCTGGATAA GCGCCTATCT

-510 TCTAATTAAA ACATTCAGAC GGGGCGGGGG ATGCGGTGGC CAAAGCACCA

TAAAACAAAA CTTCCAAGTA CTGACCAACT CACTGCAAGT TTGTGCCCCG

AGTACATCTA GGTTCAGGGG TTCTTGTCTT CATGCTCCCA ACTGCGGGCG

GATTTTTGGT CCCTTGGGAC TTTCAGTGCA GCGGCGAAGA GAGTTCTGCA

CTTGCAGGCT CCTAATGAGG GCGCAGTGGG CCTCGTGTTT CTGGTGATGC

TTCCCAGGTT GCTGGGGGCA GCAAGTGTCT CAGAGCCCAT TACTGGCTAC

ATTTTACTTC CACCAGAAAC CGAGCTGCGT CCAGATTTGC TCTCAGATGC

Figure 1 continued

GACTTGCCGC CCGGCACAGT TCCGGGGTAG TGGGGGAGTG GGCGTGGGAA

ACCGGGAAAC CCAAACCTGG TATCCAGTGG GGGGCGTGGC CGGACGCAGG

GAGTCCCCAC CCCTCCCGGT AATGACCCCG CCCCCATTCG CTAGTGTGTA

+1 (transcription start)

-10 GCCGGCGCTC TCTTTCTGCC CTGAGTCCTC AGGACCCCAA GAGAGTAAGC

TGTGTTTCCT TAGATCGCGC GGACCGCTAC CCGGCAGGAC TGAAAGCCCA

GACTGTGTCC CGCAGCCGGG ATAACCTGGC TGACCCGATT CCGCGGACAC

CGCTGCAGCC GCGGCTGGAG CCAGGGCGCC GGTGCCCCGC GCTCTCCCCG

GTCTTGCGCT GCGGGGGCGC ATACCGCCTC TGTGACTTCT TTGCGGGCCA

VRE

GGGACGGAGA AGGAGTCTGT GCCTGAGAAC TGGGCTCTGT GCCCAGCGCG

AGGTGCAGGA TGGAGAGCAA GGCGCTGCTA GCTGTCGCTC TGTGGTTCTG

CGTGGAGACC CGAGCCGCCT CTGTGGGTAA GAAGCCCACT CTTTAGTAGT

AAGGCGGAGA AGTAGGGTGC GGGCGGAGAG TGGGAATAGA AGAGGACCTA

ACTCGTAGAG CTCTAGAGAC CCTCCTCCCT TGGGTGTTCT TTCACTTACC

+490 AATGGGGAAA CTGAGGTTCA AAGACTCTTC CGAAATGACT CAGCCAGGAT

TCTACTCTCC CCCGGGCATC GGTTGGAGCG TGTCCTGCGG AGCCGTCACA

GCCCCTGGCG CTAGGTAGGC AGGAGTGGAA AGGCGGCCTG AGCCGGGGCA

GGAGATGCTC CCACTGGCAG GAACAGGCGG TCAAACGCTG GGAAGCCAGC

TCAAGCCAAG CGGCCCGGCT GGCATCAATC ACTCGGTGCT GTTGCCCACC

GCCCTAGTGG GGGGCAGGGA ATCCGCCTCT GGCTCCGCTC CCCTTTAGCT

CCAGCGTGTA AGCGCACGGA CTATGTGAGG GTAGGTCTCT TCATAGAGCA

ACACTTTCCT CCCTCAACTT TCTTTGATGC AGAATGCTAT TTTTGCTGGT

AGGAGGAAGA CGCGGCTTTC TCTTCTGTGA CAGCTTCTCC AGGTGTATTA

AACTAAATAA CTCTCCACTT ACCGACTCCA AAGCGCTGGT CCTGGGGTAA

+990 ACTCTGAAAG TCTCAGAAAC TCTTGAGCTT GGCACCTAGT TATAGGTCAC

TTTTCTTGTT TTAAAATGCC CTCTGCTTCA AGGTTAGGCC CACACTCGCT

Figure 1 continued

CTTGGGCTTT TGTGCAATAA TTTCCCTTCC CTTCCCTTCC CTTCCCTTCC
CTTCCCTTCC CTTCCCTTCC CTTCCCTTCC CTTCCCTTTC CCTCTTCCTT
TTCCTCCTCC TCTTCCTCCT CTATTTCTCT GTCATTTCCT TTTTGAAGCC
ACAGTTTGCA GATTTCCAAT CTCCACCCAT TGGAGAATGG AGAATCAGGA
AAAAAGAAGT CAATTCTGCA GAAACATTCC TTGCGCCCTA AGAGAATCGC
ATGGCTTAAA AGCATTGGCA CTGACATACG GCGCCAAGAT CGCCTGTCTA
GAGCTATTGA GTTTTCCTCA TAATGACTTG GTTCATCAGG CTAGCTCCAC
CACGAGTGCC CTCTTGTTCC TGAGAAGGCC GCACTCTCCC CCTTTCTGGG
AAGAGAAAGA CAGCCTGGAA CATGTGCTTG CCCTGGGTTC CATAGAGAAG
CAAGTTGCTT TAAAGCCCAG AGAATTCCTA GTGTAGCAGC TTAACAGCGT
CCCGTTCTCT GAATAAGATG GAGGTTGCCC TTTTGGAGTG TGTGACTTGC
XhoI (+1600)
TTAATTGGAT TGGGCTATAA TTGGTGCCAT CCAAGTCTCG AGACAGAGCC
GCTGTTGTTT TTCCTTCTGG TCTTTGAGCG GGAAGGATAA CAGTGCACAA
ATTAATTAAT GTTGGTTATC GGATTTGAAC ATAAAAGGGC TTTTATTGTA
TAGTAGCATA TGTACCTCTT GCAGTCAGAA TGAGCTGTCT AAAGAACAGA
ACCCAAACTT GCCGATGAAA ATGAATGAGG TTTAATAAAG GCGATGGATG
AGCATTAGTC ACTGATGTAA ATCTCCAGTT ATTGATAACC TCATTGACTG
GATTTGATTG CAGACATGTA TTGGTATGGG GCATCCTTTA AAGATGAGCA
+1990 TAGCCAACGT GCCTGCACTC TAAGAGAATC TATGGCTGTA TGTTATTACA
GAGACAGTTG AGAAGCTCTT AGTGGCTCTG GCGTGTAGAT CAGCGGTAGA
GCGCTGAGGC TCTGCGCTCG CTTCCTGGCA CTGAAGAATA AAGGCCATTT
ACTGTGGTGG TGCAGTGGGC GCAGTTTGTG ACGAGTTACT ACTACATTTT
CCTCACACAT CTGCCTGACT AATGAGTTCA TCAGATGAGC GTATCCAGTG
ATTGTTTGCA GGTTAATGGT TCTCAGTCAT GTTTAGAATC TACTTATCAA
ACAAATTGTT TTCTCATTTC CTGCTTCTTC TCAAACAAAG TAAGATTCCA
TTATTGAAAG GCTTGTTTAA GAGCATTTTA ACTGCTTGCC TATGTTAGGG

Figure 1 continued

ACAGTGACTT ATTTCATATT GACAAATATT ATGCCGATTA ATTGAATATG

ACTACCCAGT TCTATAGCTG TCTCAGGGCA GACCAAGAGC ATCTGTGATC

CAGTCACTTT AAATGCCATT TAAAATGCAT AATTTGTTGG TCTAGGAATA

AACACACTGT AAAGTTTAGA ATCACGGCCC AAACACAAGT CTTTAACAAT

GCCAACTAGC TTCTGAGATT CATTAATGTC ATTTAATTAC CAATGTTTTA

AAAATATGTC ATTAATTACT AAATCTATAG TTGTAACAGC AACACATGTA

CATCTTATTA AGTTGGGTAT ATTCAGGGTG GCATAGCTGT AGACTATTGC

ACATCTGTGT TGGTGAGCCA GTGGAGAACT GCCTCCTGGC TGTTCTCAGA

AGGCCACAGT GTCACGGCAT TGGCTATTTG CCTTGGCTCT TTGCTAATAC

TTTATTGACA TGGCCTCATC TTCGTTCACG TTCACTTATT TGCCCAACAA

CGTCAATGCC AGCTGAGGCC TTAGGAGTCA TCTGTTCTTA GTCAGTGCGA

ATTAGAAAGC CTGGATGCCT GCCTGCTATT AATTAGTTAT TCTTCTCTTC

+2990 TGAGACAGAG TCTCACTGTG TGGCCCAGGC TAGTCTCAAA CTTGCGGTCC

ATTTGTCTCA CTCATCAGAA TGCTGGGCTT CCAGGTGTGT GCACCACACT

AGGTAGCTCG CGTTTTAAGC TAAGAGCTGG AAGATCCTGA TGTCCTTTAC

CATGGTGGGC ATGTTACAGG TTAGTTGACT GAAAACTAGT TATCTCGCTG

TGTAATGACC TGCAGTGGTA TGTATCTCTC AAGATGCTTT TTTGCATTTC

AATCAGTTAG GTAACAAGTT CTTAAGTCTC CAGCTTGGTA TTGGCATGAG

CTCAGAGCTT TGATTAATGA GTTGGGACCC CCTAGCTATT GCTCATTAGA

CTTACACTAT TTTTAGTTTT GCTCTGAGTT TATGAATATG CATGTATGCA

TGAACTTGGG AGATATTTTT CTTCCCCAAT TCCTTTTCCT CCATTTAAAT

GTGCTGTCTT TAGAAGCCAC TGCCTCAGCT TCTGCAGCTC AGATACCAAA

GGAAGTCTGG TACACAGCAT GATAAAAGAC AATGGGACGG GGTCACAGTG

GCTCCCGTCC CTTTCAGGGG TATGGAGACG AGCTGTAGAG AGATGTCTCC

AGGGAGTTTT CATTAATCAG CAATTTAGTC AGATCTGTGC ATCCTATGCT

TTACAAGAAA TGTCAGTGGG CCTGAGATCA TCAGATGGAG GTTCATCGGG

TTTCAATGTC CCGTATCCTT TTGTAAGACC TTGAAGTTGG CAACGCAGGA

Figure 1 continued

AAACAGGAAC TCCACCCTGG TGCCGTGAAT TGCAGAGCTG TTGTGTTGGT

TTGTGACCAT CTGCCCATTC TTCCTGTTAT GACAGAGCTT GTGAACTTTA

ACTGGGACTG GGGCAAAGTC AATCCCACCT TTATACAATG AATTGCTGAA

GAGGCCTTTT AAAACTTGGA GTGTGCATTG TTTATGGAAG GGCTTTCCTA

BamHI (+3900)

TTGGATCCAA CTCTTTTCTA ATTTGTTTCT **AGGTTTGCCT GGCGATTTTC**

+3990 **TCCATCCCCC CAAGCTCAGC ACACAGAAAG ACATACTGAC AATTTTGGCA**

AATACAACCC TTCAGATTAC TTGCAGGTAA GGATTCCTTT TTGAGCCAGC

TTTCCTATGT GAAAGGACTC ATTGTTTACT GAGGTCACAA CAATTTCCAC

TATTGCAGAA GTATAATAGT ATTGTTACAA TTGTTTATAA ATCATGAGAC

TTCTAAGAAC CTATTTAATA ATGAAACAAT GGAAAAAGTC TTTTCAAACC

TTTGTACTCT TTTGCTGAGC CGTTTTCAAC ATGCACAAAC ATATTACACA

AATATAACAT ACACAGGAAC ACACATGAAT GCATGGGATG ATGTGCCTAA

AACTAGCATG TAATTGATAT TCACAATTAT TGATAAATTA GTAAAGCAAA

GGAATTCCTT ATGAATAGAG CTAAAATTCT ATCCATGTTC AAGTCACCCA

GAATGGCTTC TGGACATTTT TTTTTTTAGC TGTTTTCTAC AAGTGAAATT

CTGCCTGTAT TAGCAATTTA ATATCTAGCC AATAATATTC CTGACCATAT

GTCCTGTTCA GACCATGACC TTCATAATCT GGCTTGATGT TCTGGGCTTC

TTTCCCTCTT GCCAGCAAGA TGTCACGGTG TTGATGCTGG ATAAACTGAG

AAACAGAAGT TTTTCGCAAG AAGAGGACCT TGAATTTTGC TTTTCCCCTG

AGAGACAAGA AAGGAAACTT AGAGGAGGTG TAGCTGGGAG TGTGGTCATT

CATGAAAGAC CTGTTTGCAG GGCAGTGTGT TTTGCTGGGG ACAGTAATGA

GCCTAGATCG TAGTGCCATC CCAAGAGAGT GCTTGGTGGC AAAAAGAGCC

CTAGCAGCTT GTGGCAGTTG CCTCATATTT GAAGAATACT AAGAGGTCCC

CCGAATAACT CAGGGCTAGT GTTGATCATT GCATGTGGAG AGAATCCAAG

CCTCCTATCT AGGGTCTACA AAAGTAACCA ATGCCCAGTC TTTGGGGGAA

Figure 1 continued

+4990 AGCAAAACCA GAAAGCGATG ATAGCAGGAC CTGTTTATTT TCATTAAGTC
ATGGCATTTC CAGAGACTTT GCTCCCCCTA TTCTCAGACA CAAAGCCCAC
TTAAGATCTC CCTCTGGAGA CTGCTGGGAA CATTTCTTAA GTTCTGAAAA
AACCCTGGAG TGATTGGGCA CAGACGATCC TGTCACTTCA TGTGAGTGCT
AAGCTCTTTG GGTGATGACT CAGTGGGTCA CATTGTTTTA TTCATATTGA
CTACCTTCCG TTTGCTTTGC GGAGAATGGA AGCTATAGAA GTCTGTTTGG
TGTGGCCCTC ACAAGGCACT GTGAGCTTCT TCTCTCTGTG TGCTAACTTC
TTACTCTCCC TTGCTTATAC CCACATAGGG ACTCTGGCTT TGTTGCTGTT
CTTCAATGCT TCAGATGTGC CCTGGGTCCT GTCTGTCCTT CACACTTACT
GATGCTGCCT GGAATGCTAT TCCTCCCAAT GTGCATAGGG CCAGCTCGGT
CCAAATCCTC TCTTTTCTTT GCCTCTTTTA TATTTTCCTT CACAGTATCA
AATCACCACA GTTTATGCAA CAAACTGAAA CTTTAAAATT GTCTGTCTCC
TTATATTAGT GATAGGTTCC AGAAAGGCAC TGATTTTTTT TCTTCCCTGG
TGTACACTGG GCAACTACTC TACCACTGAG CGTGATATCC TTGGTCCCTT
AAAAGTTATC CTCTGTCCTT AATAATGCTT AGCAATCATA TTTGCTTAAA
ATATTTATTG AATGACTGCA GGAATGAATG AATGAATGAG CTAACAGAAA
ACTCATGACC ATGTGGGTGA TTTCCGAAAC AGAGTGTGAG ATCTTTGGTG
GCATGTCCTT GTAGACTGTC TGCCACCAGT ATCTATCATC TTGAAGGTGA
CTATTGAGTA GTTTATATGC ATGTGAAAAA CCAAACCTTC TATTCTCTTA
CTCATAGCCT CTCTTAATCA TAGCCCTGTG GCATGGAGTG TACCATTGAT
+5990 ATCTTCCTGG AATACTTTTT CAGGGGACAG CGGGACCTGG ACTGGCTTTG
GCCCAATGCT CAGCGTGATT CTGAGGAAAG GGTATTGGTG ACTGAATGCG
GCGGTGGTGA CAGTATCTTC TGCAAAACAC TCACCATTCC CAGGGTGGTT
GGAAATGATA CTGGAGCCTA CAAGTGCTCG TACCGGGACG TCGAC (SEQ ID NO: 1)

Figure 1 continued

Figure 2

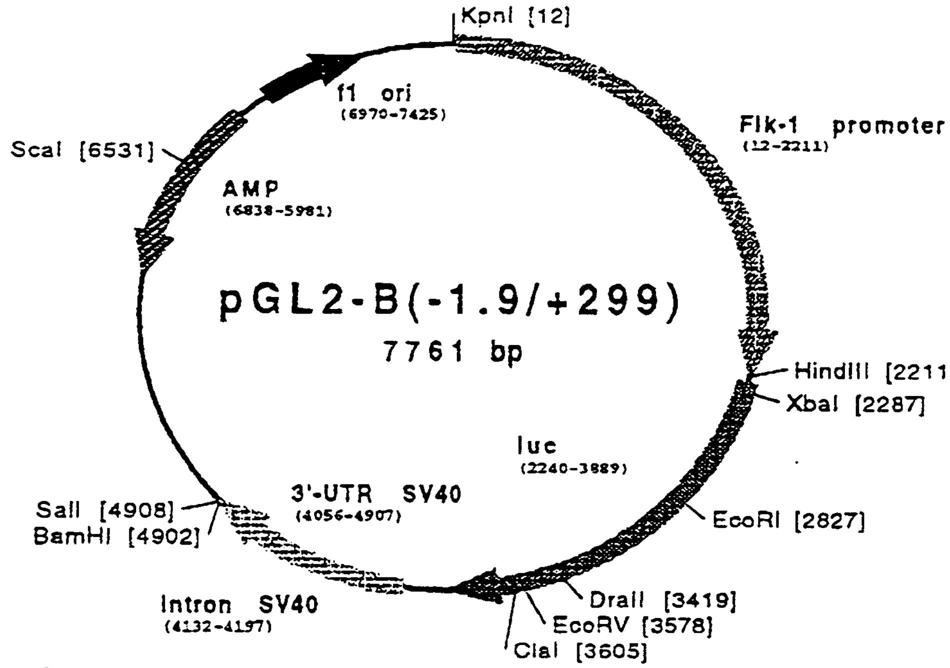

Figure 3

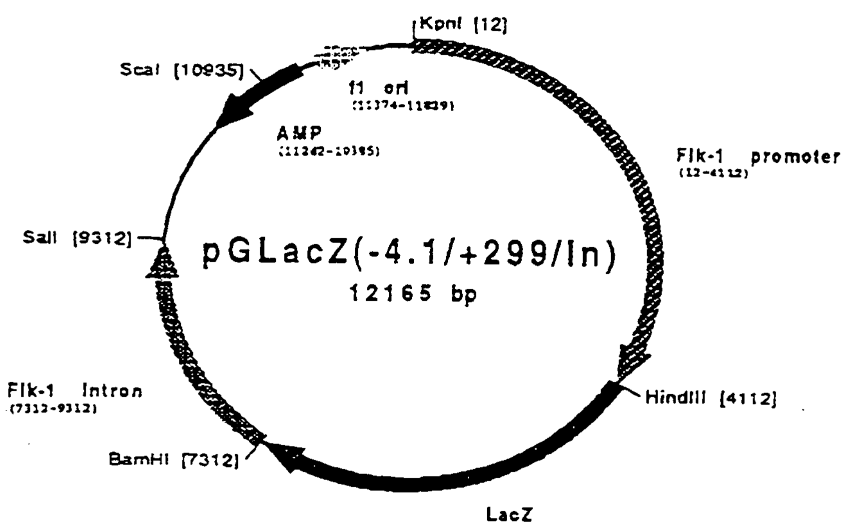

Figure 4

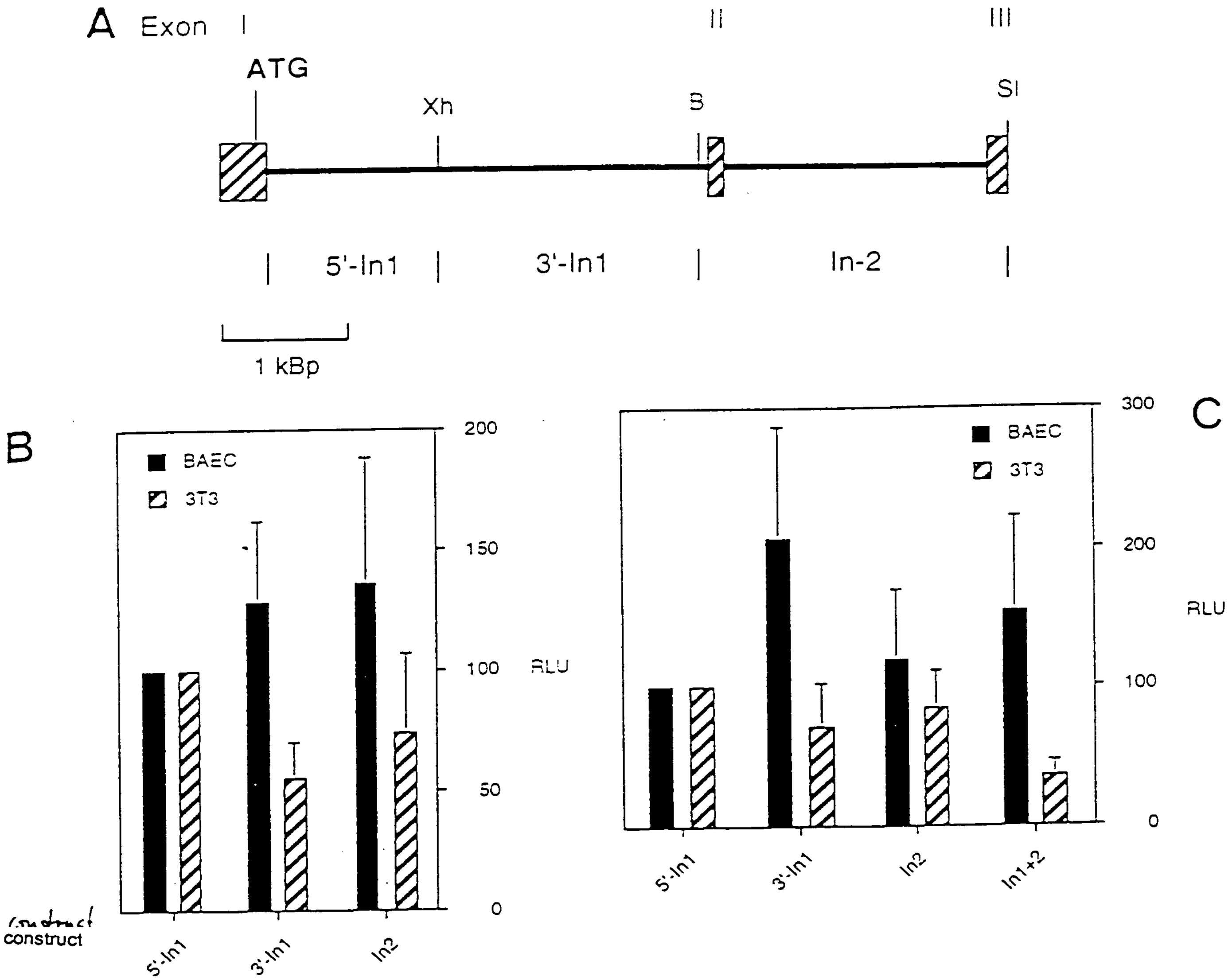
A
Exon
I
II
III
ATG
Xh
B
SI
5'-In1
3'-In1
In-2
1 kBp
B
BAEC
3T3
200
150
100
50
0
RLU
construct
construct
5'-In1
3'-In1
In2
C
BAEC
3T3
300
200
100
0
RLU
5'-In1
3'-In1
In2
In1+2

Figure 5

A

B

Figure 6

A

B

C

Figure 7

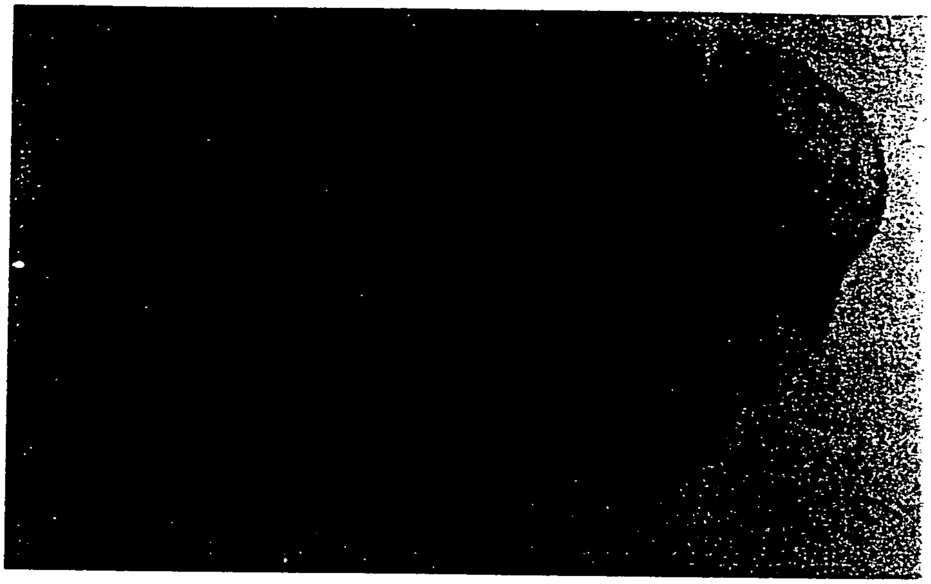

Figure 8

Figure 9

A

B

C

M 03-06-97

Figure 10